THE UNIVERSITY *of* EDINBURGH

Interference Management Techniques in Large-Scale Wireless Networks

Yi Luo

A thesis submitted for the degree of Doctor of Philosophy.
The University of Edinburgh.
December 4, 2014

Abstract

In this thesis, advanced interference management techniques are designed and evaluated for large-scale wireless networks with realistic assumptions, such as signal propagation loss, random node distribution and non-instantaneous channel state information at the transmitter (CSIT).

In the first part of the thesis, the Maddah-Ali and Tse (MAT) scheme for the 2-user and 2-antenna base station (BS) broadcast channel (BC) is generalised and optimised using the probabilistic-constrained optimisation approach. With consideration of the unknown channel entries, the proposed optimisation approach guarantees a high probability that the interference leakage power is below a certain threshold in the presence of minimum interference leakage receivers. The desired signal detectability is maximised at the same time and the closed-form solution for the receiving matrices is provided. Afterwards, the proposed optimisation approach is extended to the 3-user BC with 2-antenna BS. Simulation results show substantial sum rate gain over the MAT scheme, especially with a large spatial correlation at the receiver side.

In the second part, the MAT scheme is extended to the time-correlated channels in three scenarios, in which degrees of freedom (DoF) regions as well as achievability schemes are studied: 1) 2-user interference channel (IC) using imperfect current and imperfect delayed CSIT; 2) K-user BC with K-antenna BS using imperfect current and perfect delayed CSIT; 3) 3-user BC with 2-antenna BS using imperfect current and perfect delayed CSIT. Notably, the consistency of the proposed DoF regions with the MAT scheme and the ZF beamforming schemes using perfect current CSIT consents to the optimality of the proposed achievability schemes.

In the third part, the performance of the ZF receiver is evaluated in Poisson distributed wireless networks. Simple static networks as well as dynamic networks are studied. For the static network, transmission capacity is derived whereby the receiver can eliminate interference from nearby transmitters. It is shown that more spatial receive degrees of freedom (SRDoF) should be allocated to decode the desired symbol in the presence of low transmitter intensity. For the dynamic network, in which the data traffic is modelled by queueing theory, interference alignment (IA) beamforming is considered and implemented sequentially. Interestingly, transmitting *one* data stream achieves the highest area spectrum efficiency.

Finally, a distance-dependent IA beamforming scheme is designed for a generic 2-tier heterogeneous wireless network. Second-tier transmitters partially align their interferences to the dominant cross-tier interference overheard by the receivers in the same cluster. Essentially, the proposed IA scheme compromises between enhancing the signal-to-interference ratio and increasing the multiplexing gain. It is shown that acquiring accurate distance knowledge brings insignificant throughput gain compared to statistical distance knowledge. Simulation results validate the derived expressions of success probabilities as well as throughput, and show that the distance-dependent IA scheme significantly outperforms the traditional IA scheme in the presence of path-loss effect.

Lay Summary

Modern wireless networks are providing better reliability and a higher data rate through more intensive infrastructure deployment. As more and more devices and infrastructures are activated, interference management techniques are facing a greater challenge, for two reasons. First, interference is becoming more heterogeneous because wireless network infrastructures are becoming diverse and, secondly, as wireless networks are becoming larger-scaled and data traffic is becoming uncertain, it is difficult to determine how to model the interference accurately. Limited spectrum resource exacerbates the problem, because it is essential to reduce interference in order to improve services further.

In the first half of the thesis, delayed channel state information at transmitter (CSIT) is exploited to mitigate interference. Exploiting delayed CSIT is appealing in practice because it can be easily obtained. More specifically, the achievable degrees-of-freedom (DoF) regions, as well as the corresponding achievability schemes, are investigated using the delayed CSIT in the time-correlated channels, where DoF is an effective metric for the data rate when the signal-to-noise ratio (SNR) is high. Further, the Maddah-Ali and Tse (MAT) scheme (known as the first scheme to exploit delayed CSIT) is generalised and optimised using a probabilistic-constrained optimisation approach. The improved scheme outperforms the original MAT scheme significantly when the channel correlation is large.

In the second half of the thesis, interference in large-scale randomly distributed wireless networks is modelled and analysed using an emerging technique named stochastic geometry. The performance of a zero-forcing (ZF) receiver, which is a classic and tractable receiver structure, is analysed in such wireless networks and also in asynchronous wireless networks. The trade-off between increasing the desired signal power and reducing the interference power is investigated. Moreover, a distance-dependent interference alignment (IA) scheme is tailored for heterogeneous wireless networks in the presence of path-loss effect. It is shown that the proposed IA scheme substantially outperforms the original scheme, which ignores the path-loss effect. Finally, numerical results validate the effectiveness of stochastic geometry and show that interference can be accurately modelled.

Declaration of Originality

I hereby declare that the research recorded in this thesis and the thesis itself was composed and originated entirely by myself in the School of Engineering at The University of Edinburgh.

Yi Luo

Acknowledgements

"To be a better man."

First of all, I would like to give my gratitude to my supervisor, Dr Tharmalingam Ratnarajah, who provided me the opportunity to conduct this interesting and meaningful research and also spent precious time to guide me. The valuable experience during the past three years would be extremely rewarding for my future career. Meanwhile, I would like to appreciate all the constructive discussions with my colleagues, as well as their careful proof-reading and support beyond academia. I could never accomplish this thesis work without the support from HIATUS funding and sincerely hope that my efforts will repay the expectations.

I would also present my most sincere respect to my parents, who always support whatever my decision is and let me understand that there is a solid foundation waiting for me back there in China. I understand that what I need to thank for is far beyond past three years, which dates back to all the picking-ups from kindergarten and all the breakfasts in the mornings of high school. Meanwhile, I would like to give special thanks to Miss Xu for the considerate care during the past two and half years. Finally, thank all the other family members. The most valuable treasure I have is your love and support.

Last but not least, I want to encourage myself whenever sense of loss sweeps: always try your best and always be thankful.

Contents

Declaration of Originality . . . iii
Acknowledgments . . . iv
List of Figures . . . ix
List of Tables . . . xi
Acronyms and Abbreviations . . . xii
Nomenclature . . . xiv

1 Introduction **1**
1.1 Motivations . . . 1
1.2 Thesis Layout and Contributions . . . 3

2 Background **6**
2.1 Wireless Channel . . . 6
2.1.1 Fading . . . 6
2.1.2 Path-Loss Effect . . . 7
2.2 MIMO System . . . 8
2.2.1 CSIT and Beamforming Matrix Acquisition . . . 9
2.2.2 Broadcast Channel and Interference Channel . . . 10
2.2.3 Interference Cancellation . . . 11
2.2.4 Degrees of Freedom Region . . . 12
2.2.5 A Brief Introduction to Interference Alignment . . . 13
2.3 Stochastic Geometry . . . 15
2.3.1 Wireless Networks and Point Process . . . 15
2.3.2 Borel σ-algebra . . . 16
2.3.3 Measurability in 2-Dimensional Space . . . 17
2.3.4 Fundamentals of Point Process . . . 19
2.3.5 Poisson Point Process . . . 19
2.3.6 Poisson Clustered Process . . . 20
2.3.7 Poisson Clustered Process and Heterogeneous Wireless Networks . . . 21
2.3.8 Voronoi Cell . . . 22

2.3.9 Outage Probability . . . 23
2.3.10 Throughput . . . 23
2.3.11 Transmission Capacity . . . 24

3 Probabilistic-Constrained Optimisation for MISO BC Using Delayed CSIT 25

3.1 Introduction . . . 25
3.2 System Model . . . 27
3.2.1 The MAT Scheme for $\mathcal{S}_{MAT}^{(1)}$ . . . 27
3.2.2 The MAT Scheme for $\mathcal{S}_{MAT}^{(3)}$ . . . 29
3.3 Probabilistic-Constrained Optimisation for $\mathcal{S}_{MAT}^{(1)}$ . . . 31
3.4 Probabilistic-Constrained Optimisation for $\mathcal{S}_{MAT}^{(3)}$ . . . 36
3.5 Simulation Results . . . 39
3.6 Summary . . . 43

4 Degrees of Freedom Regions Using Delayed CSIT in The Time-Correlated Channels 44

4.1 Introduction . . . 44
4.2 System Model . . . 45
4.3 Scheme for $\mathcal{S}_{MAT}^{(1)}$ with Imperfect Current and Delayed CSIT . . . 46
4.3.1 Degrees of Freedom Region . . . 47
4.3.2 Scheme $\mathcal{X}_1$ for Vertex $M^{(1)}$ . . . 47
4.3.3 Scheme $\mathcal{X}_2$ for Vertex $A^{(1)}$ . . . 54
4.4 Improved Scheme for $\mathcal{S}_{MAT}^{(1)}$ with Imperfect Current and Delayed CSIT . . . 55
4.4.1 Degrees of Freedom Region . . . 55
4.4.2 Scheme $\mathcal{X}_3$ for Vertices $M_1^{(2)}$ and $M_2^{(2)}$ . . . 57
4.4.3 Scheme $\mathcal{X}_4$ for Vertices $A_1^{(2)}$ and $A_2^{(2)}$ . . . 63
4.5 Setting $\mathcal{S}_{MAT}^{(2)}$ with Imperfect Current and Perfect Delayed CSIT . . . 64
4.5.1 Degrees of Freedom Region for 3-User Case . . . 65
4.5.2 Achievability Scheme $\mathcal{X}_5$ for Vertex $M^{(3)}$ . . . 66
4.5.3 Achievability Scheme $\mathcal{X}_6$ for Vertices $D_1^{(3)}$, $D_2^{(3)}$ and $D_3^{(3)}$ . . . 70
4.5.4 Extension to the K-User Case . . . 71
4.6 Setting $\mathcal{S}_{MAT}^{(3)}$ with Imperfect Current and Perfect Delayed CSIT . . . 73
4.6.1 Degrees of Freedom Region . . . 74
4.6.2 Scheme $\mathcal{X}_9$ for Vertex $M^{(4)}$ . . . 77
4.7 Simulation Results . . . 81
4.8 Summary . . . 83

5 Interference Cancellation in Poisson Distributed Wireless Networks **85**

5.1 Introduction . . . 85

5.2 System Model . . . 87

5.3 ZF Receiver in A Static Network . . . 88

5.3.1 Nearest Neighbour Distance . . . 89

5.3.2 Average Signal-to-Interference-Noise Ratio . . . 89

5.3.3 Transmission Capacity with Rayleigh Fading . . . 90

5.4 Influence of the Feasibility Condition of the ZF Receiver . . . 92

5.4.1 Feasibility Condition . . . 92

5.4.2 Perfect ZF Condition . . . 93

5.4.3 Imperfect ZF Condition . . . 95

5.5 ZF Receiver with IA Beamforming in A Dynamic Network . . . 97

5.5.1 A More Specific System Model . . . 97

5.5.2 Interference Alignment Implementation . . . 98

5.5.3 Outage Probability and Transmission Capacity . . . 103

5.6 Simulation Results . . . 105

5.7 Summary . . . 110

6 Throughput for A Generic Heterogeneous Wireless Network Using Interference Alignment **112**

6.1 Introduction . . . 112

6.1.1 Contributions . . . 113

6.1.2 Organisations . . . 114

6.2 System Model . . . 115

6.3 Interference Mitigation . . . 119

6.3.1 Nearest Cross-Tier Interference Decomposition . . . 120

6.3.2 Feasibility Condition of Interference Alignment . . . 121

6.3.3 Distance-Dependent Interference Alignment . . . 123

6.4 Outage Probability and Throughput . . . 127

6.4.1 Remaining Cross-Tier Interference . . . 128

6.4.2 Inter-Cluster Interference . . . 129

6.5 Simulation Results . . . 131

6.6 Summary . . . 136

7 Conclusions and Future Work **138**

7.1 Conclusions 138
7.2 Limitations and Future Work 141
7.2.1 Further Exploit the Delayed CSIT 141
7.2.2 Realise Full Potential of Stochastic Geometry 142
7.2.3 Other Future Directions 143

Appendix A Proof of Theorem 1 **144**

Appendix B Proof of Theorem 2 **146**

Appendix C Proof of Theorem 8 **148**

Appendix D Proof of Theorem 9 **150**

Appendix E Proof of Theorem 10 **151**

Appendix F List of Publications **154**
F.1 Accepted Publications 154
F.2 Papers Under Revision 155

List of Figures

2.1 A point-to-point MIMO wireless communication system. 9
2.2 Broadcast channel and interference channel. 10
2.3 Sum rates of the ZF scheme, the TDMA scheme and the greedy beamforming scheme. 12
2.4 Degrees-of-freedom region of a 2-user MISO BC. 13
2.5 An example of MIMO IA for a 3-user IC. 16
2.6 Locations of BSs near River Thames, central London. 17
2.7 A poisson point process with unit intensity. 20
2.8 A Matérn process in the Euclidean plane. 21
2.9 The BSs and terminals in a heterogeneous cellular network. 22
2.10 Voronoi tessellation of a PPP in 2-dimensional space. 23

3.1 The MAT scheme for the 2-user MISO BC. 28
3.2 Sum rate for 2-user MISO BC versus γ with various SNRs. 40
3.3 Sum rate of 2-user MISO BC with a large spatial correlation at the receiver side. . 41
3.4 Sum rate of 2-user MISO BC versus the spatial correlation at the receiver side. . . 42

4.1 The DoF region of 2-user MISO IC using the basic scheme. 48
4.2 The DoF region of 2-user MIMO IC using the advanced scheme. 56
4.3 The DoF region of 3-user BC with *three* BS antennas. 65
4.4 The DoF region of 3-user MIMO BC with *two* BS antennas. 75
4.5 Sum rate achieved by Scheme $\mathcal{X}_1$ compared with the TDMA scheme and the ZF scheme. 82
4.6 Sum rate achieved by Scheme $\mathcal{X}_3$ compared with the TDMA scheme and the ZF scheme. 83

5.1 Dynamic data traffic around R_0 in n time slots, $n \geq 3$. 99
5.2 Average signal-to-noise ratio of a ZF receiver with various coverage factors. . . . 106
5.3 Access probability versus μ_h in early stage and stable stage. 107
5.4 Analysis and simulation of dimension of the interference-free subspace. 108

5.5 Analysis and simulation of the success probability versus r_z. 108
5.6 Analysis and simulation of the transmission capacity versus outage probability. . 109

6.1 The distribution of transmitters in tier-A and tier-B. 115
6.2 Voronoi cells of transmitters in tier-A with Poisson intensity $\lambda_A = 0.002$. 118
6.3 Signal space at R_0 when interference are maximum overlapped. 121
6.4 Number of data streams from L_0 that should be eliminated. 132
6.5 The throughput only affected by the nearest transmitter in tier-A. 133
6.6 The success probability at R_0 affected by the remaining tier-A transmitters. . . . 134
6.7 The success probability at R_0 affected by the inter-cluster transmitters. 135
6.8 The overall expected throughput versus tier-A transmit power. 135

List of Tables

2.1 Path-loss exponent values in various environments [1]. 8

3.1 Symbols received in all of the eight time slots of the 3-user MAT scheme with 2-antenna BS. 30

4.1 Private symbol vector contained in each of the common symbol vector. 78

6.1 Default values for parameters. 131

Acronyms and Abbreviations

BC	Broadcast Channel
BS	Base Station
CDF	Cumulative Density Function
CSI	Channel Statement Information
CSIT	Channel Statement Information at the Transmitter
CSMA	Carrier Sense Multiple Access
DoF	Degrees of Freedom
HCPP	Hard Core Point Process
IA	Interference Alignment
IC	Interference Channel
IBC	Interference Broadcast Channel
i.i.d.	Independent and Identical Distribution
LTE-A	Long Term Evolution-Advanced
MAC	Multiple Access Control
MANET	Mobile Ad-Hoc Network
MAT	Maddah-Ali and Tse
MGF	Moment Generating Function
MIMO	Multiple-Input-Multiple-Output
MISO	Multiple-Input-Single-Output
MMSE	Minimum Mean Square Error
MRT	Maximum Ratio Transmission
NFC	Near Field Communication
OFDM	Orthogonal Frequency-Division Multiplexing
PCP	Poisson Clustered Process
PDF	Probability Density Function
PPP	Poisson Point Process

Rx	Receiver
SNR	Signal-to-Noise Ratio
SIC	Successive Interference Cancellation
SINR	Signal-to-Interference-Noise Ratio
SIR	Signal-to-Interference Ratio
SISO	Single-Input-Single-Output
SRDoF	Spatial Receive Degrees of Freedom
TDMA	Time Division Multiple Access
Tx	Transmitter
ZF	Zero-Forcing

Nomenclature

$\bigcap$	Union operation
$\succcurlyeq \mathbf{0}$	Positive semi-definite
$f(x) \doteq g(x)$	$\lim_{P\to\infty} \frac{f(x)}{g(x)}$ is a constant not scaled of P
$\approx$	Approximately equal to
$\propto$	Proportional to
$\emptyset$	Null set
$\mathbb{C}$	Complex number
$\mathbb{C}^{m\times n}$	$m \times n$ complex matrix
$\mathbb{N}$	Natural number
$\mathbb{R}$	Real number
$\mathbb{R}^d$	d-dimensional Euclidean space
$\mathbb{R}_+^d$	d-dimensional array with positive components
$\mathbb{Z}$	Integer
α	Power exponent of estimation error of the current CSIT
β	Power exponent of estimation error of the delayed CSIT
χ_k^2	Chi-squared distribution with k degrees of freedom
Φ	Point process
Φ_A	Point process of tier-A transmitters
Φ_B	Point process of tier-B transmitters
Φ_P	Parent point process
Φ_T	Point process of transmitters
$\bar{\phi}_i$	Untranslated point process of cluster i
ϕ_i	Translated point process of cluster i
ε	Standard error of channel estimation error
ϵ	Outage probability
ϵ_p	Outage probability under perfect zero-forcing condition
ϵ_{ip}	Outage probability under imperfect zero-forcing condition
ϵ_{lb}	Lower bound of outage probability
ϵ_{ub}	Upper bound of outage probability
ϵ_p^{lb}	Lower bound of outage probability under perfect zero-forcing condition

ϵ_p^{ub}	Upper bound of outage probability under perfect zero-forcing condition
θ	Pre-specified signal-to-interference ratio threshold
ς	Path-loss exponent
λ	Transmitter intensity in a Poisson network
λ_∞	Transmitter intensity in stable stage
λ^ϵ	Maximum transmitter intensity to keep the outage probability ϵ
λ_A	Tier-A transmitter intensity
λ_B	Tier-B parent point intensity
σ_h^2	Variance of real and imaginary parts of channel entries
$\sigma_{\breve{\eta}_i}^2$	Variance of interference reconstruction error
σ_D^2	Required variance of digitisation error
γ	Parameter for the allowed interference leakage power
$\Upsilon_i(\mathbf{w}_i)$ or $\bar{\Upsilon}_i(\bar{\mathbf{w}}_i)$	Numerator containing $\mathbf{w}_i$ only
$\Psi_i(\mathbf{w}_i)$ or $\bar{\Psi}_i(\bar{\mathbf{w}}_i)$	Denominator containing $\mathbf{w}_i$ only
η_i	Overheard interference at Rx-i
$\breve{\eta}_i$	Reconstructable part of η_i
$\bar{\breve{\eta}}_i$	Digitised version of $\breve{\eta}_i$
$\ddot{\eta}_i$	Unreconstructable part of η_i
$\tilde{\eta}_i$	Digitisation error when digitising $\breve{\eta}_i$
$\boldsymbol{\eta}_i$	Overheard interference vector at Rx-i
$\boldsymbol{\eta}_i(s,t)$	Overheard interference vector at Rx-i at time slot t of Phase s
$\breve{\boldsymbol{\eta}}_i(s,t)$	Reconstructable part of $\boldsymbol{\eta}_i(s,t)$
$\bar{\breve{\boldsymbol{\eta}}}_i(s,t)$	Digitised version of $\boldsymbol{\eta}_i(s,t)$
$\ddot{\boldsymbol{\eta}}_i(s,t)$	Unreconstructable part of $\boldsymbol{\eta}_i(s,t)$
$\tilde{\boldsymbol{\eta}}_i(s,t)$	Digitisation error when digitising $\breve{\boldsymbol{\eta}}_i(s,t)$
μ_h	Parameter in the cumulative distribution function of holding time
μ_v	Arrival rate
τ_i	Arrival time of T_i
$\boldsymbol{\Sigma}_i$	Arbitrary i columns of the $\bar{d}_A \times \bar{d}_A$ identity matrix
$\tilde{\boldsymbol{\Sigma}}_i$	The complementary columns of $\boldsymbol{\Sigma}_i$ in the identity matrix
$\mathcal{A}$	σ-algebra on subset of A
$\mathcal{A}_c$	Channel state information area
$\mathcal{A}_l$	Area outside the nearest transmitter
$\mathcal{A}_z$	Zero-forcing area

$\mathcal{B}^2$	2-dimensional Borel set
c	Parameter of remaining holding time of dominant transmitters
c_i	Common symbol with index i
$\mathbf{c}_i$ or $\bar{\mathbf{c}}_i$	Common symbol vector with index i
$\mathbf{c}_{ij}$ or $\bar{\mathbf{c}}_{ij}$	Order-2 common symbol vector with index ij
$\mathbf{c}_{123}^{(i)}$	The i-th order-3 common symbol vector
$\mathbf{c}_i(s,t)$	Common symbol vector at time slot t of Phase s with index i
$\mathcal{C}$	Transmission capacity
$\mathcal{C}_p^{lb}$	Lower bound of transmission capacity under perfect ZF condition
$\mathcal{C}_p^{ub}$	Upper bound of transmission capacity under perfect ZF condition
$\mathcal{C}_{ip}^{lb}$	Lower bound of transmission capacity under imperfect ZF condition
$\mathcal{C}_{ip}^{ub}$	Upper bound of transmission capacity under imperfect ZF condition
$\mathcal{C}_{lb}$	Lower bound of transmission capacity
$\mathcal{C}_{ub}$	Upper bound of transmission capacity
$\mathcal{CN}(\cdot,\cdot)$	Complex normal distribution
$\bar{d}$	Number of data streams
$\bar{d}_a$	Number of data streams being treated as interference
$\hat{d}_a$	Optimal number of data streams being treated as interference
$\bar{d}_A$	Number of data streams from tier-A transmitters
$\bar{d}_B$	Number of data streams from tier-B transmitters
$\bar{d}_I$	Dimension of the interference subspace
d_i	Degrees-of-freedom achieved by Rx-i
$\mathbf{f}_q$	Decorrelator at R_0 for the q-th data stream
g_i or $\bar{g}_i$	Parameter in the generalised receiving matrix for the MAT scheme
$\hat{g}_i^2$	$\max\{1, \lvert g_i\rvert^2\}$
$\mathbf{G}_{ij}$	The effective MIMO channel in the 2-user MAT scheme
h	Fading coefficient
$h_{ij,1}(t)$	Channel entry from the first antenna of Tx-j to Rx-i at time slot t
$h_{ij,1}$	Channel entry from the first antenna of Tx-j to Rx-i
$h_{i,1}(t)$	Channel entry from the first antenna of the BS to Rx-i at time slot t
$\mathbf{h}_{ij}$	Channel vector from Tx-j to Rx-i
$\hat{\mathbf{h}}_{ij}$	Current estimate of $\mathbf{h}_{ij}$
$\mathbf{h}_{ij}(t)$	Channel vector from Tx-j to Rx-i at time slot t
$\hat{\mathbf{h}}_{ij}(t)$	Current estimate of $\mathbf{h}_{ij}(t)$

$\tilde{\mathbf{h}}_{ij}(t)$	Current estimation error of $\mathbf{h}_{ij}(t)$
$\check{\mathbf{h}}_{ij}(t)$	Delayed estimate of $\mathbf{h}_{ij}(t)$
$\ddot{\mathbf{h}}_{ij}(t)$	Delayed estimation error of $\mathbf{h}_{ij}(t)$
$\mathbf{h}_i(t)$	Channel vector from the base station to Rx-i at time slot t
$\mathbf{H}_i$	Channel matrix from the base station to Rx-i
$\hat{\mathbf{H}}_i$	Current estimate of $\mathbf{H}_i$
$\mathbf{H}_{i,1}$	Left squared part of $\mathbf{H}_i$
$\mathbf{H}_i(t)$	Channel matrix from the base station to Rx-i at time slot t
$\hat{\mathbf{H}}_i(t)$	Current estimate of $\mathbf{H}_i(t)$
$\mathbf{H}_{i,1}(t)$	Left squared part of $\mathbf{H}_i(t)$
$\mathbf{H}_{ij}$	Channel matrix from Tx-j to Rx-i
$\mathbf{H}_{ij,1}$	Left squared part of $\mathbf{H}_{ij}$
$\hat{\mathbf{H}}_{ij}$	Current estimate of $\mathbf{H}_{ij}$
$\mathbf{H}_{iL_j}$	Channel matrix from L_j to R_i
$\hat{\mathbf{H}}_{iL_j}$	Current estimate of $\mathbf{H}_{iL_j}$
$\tilde{\mathbf{H}}_{iL_j}$	Current estimation error of $\mathbf{H}_{iL_j}$
$\mathbf{H}_{ij}(s,t)$	Channel matrix from Tx-j to Rx-i at time slot t of Phase s
$\hat{\mathbf{H}}_{ij}(s,t)$	Current estimate of $\mathbf{H}_{ij}(s,t)$
$\check{\mathbf{H}}_{ij}(s,t)$	Delayed estimate of $\mathbf{H}_{ij}(s,t)$
$\ddot{\mathbf{H}}_{ij}(s,t)$	Delayed estimation error of $\mathbf{H}_{ij}(s,t)$
$\mathbf{H}_{ij,1}(s,t)$	Left square part of $\mathbf{H}_{ij}(s,t)$
$\mathfrak{H}$	All the channel state information at Rx-1 when decoding symbols
I_q	Total interference at the q-th data stream
I_c	Cross-tier interference
I_A	Remaining cross-tier interference
I_n	Nearest cross-tier interference
$I_n^{(1)}$	Nearest cross-tier interference treated as interference
$I_n^{(2)}$	Nearest cross-tier interference treated as shot noise
I_{inter}	Inter-cluster interference
$\mathbf{I}^{(k)}$	The $k \times k$ identity matrix
$\hat{k}$	Coverage factor
$\hat{k}_0$	Number of dominant transmitters
K	Number of users in a system or cluster
l_{ij}	Distance from T_j to R_i

l_{iL_j}	Distance from L_j to R_i
l_k	Distance to the k-th nearest transmitter
L	Edge length of the simulation plane
L_i	Transmitter i in tier-A
M	Number of transmitter antennas
$n_i(t)$	Additive white noise at Rx-i at time slot t
$\mathbf{n}_i$	Additive white noise vector at Rx-i
$\mathbf{n}_i(t)$	Additive white noise vector at Rx-i at time slot t
$\mathbf{n}_i(s,t)$	Additive white noise vector at Rx-i at time slot t of Phase s
N	Number of receiver antennas
N_A	Number of antennas at tier-A transmitters
N_B	Number of antennas at tier-B transmitters and receivers
$\mathcal{N}$	Normal distribution
$\mathbf{N}_i$	Covariance matrix of the white noise vector at Rx-i
o	The origin of a plane
p_s	Success probability
p_s^n	Success probability affected by the nearest cross-tier interference
p_s^A	Success probability affected by the remaining cross-tier interference
p_s^{inter}	Success probability affected by the inter-cluster interference
P	Signal-to-noise ratio
P_A	Transmit power of tier-A transmitters
P_r	Received power in Watt
P_t	Transmit power in Watt
$\mathbf{q}_{ij}^{\perp}$	Normalised vector lying in the orthogonal space of $\hat{\mathbf{h}}_{ij}$
$\mathbf{q}_{ij}(t)$	Normalised vector lying in the range of $\hat{\mathbf{h}}_{ij}(t)$
$\mathbf{q}_{ij}^{\perp}(t)$	Normalised vector lying in the orthogonal space of $\hat{\mathbf{h}}_{ij}(t)$
$\mathbf{Q}_{ij}^{\perp}$	Matrix lines in the orthogonal space of vectors in $\hat{\mathbf{H}}_{ij}$
$\mathbf{Q}_{ij}(s,t)$	Matrix lines in the range of $\hat{\mathbf{H}}_{ij}(s,t)$
$\mathbf{Q}_{ij}^{\perp}(s,t)$	Matrix lines in the orthogonal space of vectors in $\hat{\mathbf{H}}_{ij}(s,t)$
$\mathbf{Q}_i(t)$	Matrix lines in the range of $\hat{\mathbf{H}}_i(t)$
$\mathbf{Q}_i^{\perp}(t)$	Matrix lines in the orthogonal space of vectors in $\hat{\mathbf{H}}_i(t)$
$\mathbf{Q}_{i,j}^{\perp}(t)$	Matrix lines in the orthogonal space of vectors in $\hat{\mathbf{H}}_i(t)$ and $\hat{\mathbf{H}}_j(t)$
$\mathbf{Q}_{i,j}^{\perp}$	Matrix lines in the orthogonal space of vectors in $\hat{\mathbf{H}}_i$ and $\hat{\mathbf{H}}_j$
r_a	Radius of a cluster

r_c	Radius of channel state information area
r_r	Spatial correlation coefficient at the receiver side
r_z	Radius of zero-forcing area
r_∞	Distance to the furthest interferer
R	Data rate in one data stream
R_i	Rx-i
$\mathbf{R}_r$	Receiver correlation matrix
$s_{0,q}$	Desired signal power at the q-th data stream of Rx-0
$s_i^{(j)}$	Private symbol for Rx-i
$\mathbf{s}_{L_i}$	Symbol vector sent by L_i
$\mathbf{s}_i(s,t)$	Private symbol vector for Rx-i at time slot t of Phase s
$\mathbf{s}_i^{(j)}$ or $\bar{\mathbf{s}}_i^{(j)}$	Private symbol vector for Rx-i
$\mathbf{s}_i$ or $\bar{\mathbf{s}}_i$	Private symbol vector for Rx-i
S	Total number of phases
$\mathcal{S}_{MAT}^{(1)}$	Setting with 2-user and 2-antenna transmitter(s)
$\mathcal{S}_{MAT}^{(2)}$	Setting with K-user and K-antenna transmitter(s)
$\mathcal{S}_{MAT}^{(3)}$	Setting with 3-user and 2-antenna transmitter(s)
T_i	Tx-i
T_f	Transmitter selected as the reference one
$\hat{T}_s$	Number of time slots in Phase s
$\mathcal{T}$	Throughput
$\mathbf{U}_i$	Receiving matrix at Rx-i
$\mathbf{V}_i$	Beamforming matrix at Tx-i
$\mathbf{V}_{L_i}$	Beamforming matrix at L_i
$\mathbf{w}_i$ or $\bar{\mathbf{w}}_i$	Generalised beamforming vector in the MAT scheme
$\mathbf{x}$	Beamformed symbol vector at the base station
$\mathbf{x}(t)$	Beamformed symbol vector at the base station at time slot t
$\mathbf{x}_i$	Beamformed symbol vector at Tx-i
$\mathbf{x}_i(t)$	Beamformed symbol vector at Tx-i at time slot t
$\mathbf{x}_i(s,t)$	Beamformed symbol vector at Tx-i at time slot t of Phase s
$\bar{\mathbf{x}}_i(s,t)$	Private symbol part of $\mathbf{x}_i(s,t)$
$y_i(t)$	Received signal at Rx-i at time slot t
$\mathbf{y}_i$	Received signal at Rx-i
$\hat{\mathbf{y}}_i$	Post-processing received signal at Rx-i

$\mathbf{y}_i(t)$	Received signal at Rx-i at time slot t
$\mathbf{y}_i(s,t)$	Received signal at Rx-i at time slot t of Phase s
$(x)^+$	$\max\{0,x\}$
$(\cdot)'$	Derivative operation
$(\cdot)^H$	Conjugation transpose
$(\cdot)^T$	Transpose
$(\cdot)^*$	Conjugation
$(\cdot)^\perp$	Orthogonal vector or space
$\lvert\cdot\rvert$	Lebesgue measure for a set or absolute value for a number
$\lVert\cdot\rVert$	2-norm or Euclidean norm
$\lVert\cdot\rVert_F$	Frobenius Norm
$(\cdot)!$	Factorial
$(\cdot)!!$	Double factorial
$\Lambda(\cdot)$	Intensity measure
$\Gamma(\cdot)$	Gamma function
$\Gamma(\cdot,\cdot)$	Upper incomplete gamma function
$\gamma(\cdot,\cdot)$	Lower incomplete gamma function
$\mathrm{erf}(\cdot)$	Error Function
$\delta_x(\cdot)$	Dirac measure
$\Phi(\cdot)$	Counting measure of point process Φ
$\mathbf{1}_A(\cdot)$	Indicator function of the set A
${}_2F_1[\cdot]$	Hypergeometric function
$b(x,r)$	A 2-dimensional ball centred at x with radius r
$C(\cdot)$	Channel capacity
$\bar{C}(\cdot)$	Capacity region of a system
$\bar{d}_0(\cdot)$	Dimension of interference-free space at R_0
$\hat{d}_0(\cdot)$	SRDoF that can be utilised to decode any data stream
$\mathbb{E}[\cdot]$	Expectation operation
$f^{-1}(\cdot)$	Inverse function
$f_{L_0}(\cdot)$	Probability density function of the distance from L_0 to R_0
$f_B(\cdot)$	Radius intensity of the transmitters in a tier-B cluster
$f_n(\cdot,\cdot)$	Feasibility condition of the zero-forcing receiver
$f_{\chi^2}(\cdot,\cdot)$	Probability density function of chi-squared distribution
$f_\Omega(\cdot)$	Radius intensity function of the transmitters

$\tilde{F}_{\Gamma}(\cdot)$	Complementary cumulative distribution function of Gamma distribution
$F_h(\cdot)$	Cumulative distribution function of the holding time
$F_v(\cdot)$	Cumulative distribution function of the arrival interval
$F_k(\cdot)$	Cumulative distribution function of the distance to the k-th nearest transmitter
$\mathfrak{G}_{PPP}(\cdot)$	Probability generating function of PPP
$\mathfrak{G}_{PCP}(\cdot)$	Probability generating function of PCP
$\mathcal{G}(\cdot)$	Meijer-G function
$I(\cdot;\cdot)$	Mutual information
$\log(\cdot)$	Logarithmic function with base 2
$\ln(\cdot)$	Natural logarithmic function
$\mathcal{L}_I(s)$	Laplace transform of interference I
$M_I(t)$	Moment generating function of interference I
$\Pr(\cdot)$	Probability that an event occurs
$V(\cdot)$	Voronoi cell of a point
$\mathbf{R}(\cdot)$	Data rate tuple
$\mathrm{span}(\cdot)$	Linear span

Chapter 1
Introduction

Since the release of the first-generation (1G) standard for wireless mobile phones during the 1980s, modern wireless communication networks have gone through more than 30 years of continuous evolution. During this time, the world has witnessed an explosive growth of wireless communication services, and mobile phones have evolved from an analogue standard to the newly released Long-Term Evolution-Advanced (LTE-A) standard. Along with this evolution, the reliability and data rates of mobile communication networks have increased tremendously. This trend is widely believed to continue in the coming decades with the advent of smart phones and Internet-based services such as cloud storage and online video streaming. In order to meet the expectations for future wireless networks, various kinds of infrastructure have been deployed, such as the femto base station (BS), the pico BS and the macro BS. At the same time, near-field communication (NFC) and wearable devices such as smart watches have started to emerge. Therefore, devices in various kinds of networks such as mobile ad-hoc networks (MANETs), personal area networks (PANs) and cellular networks will merge with devices in other networks [2]. Together with limited spectrum resources, managing interference is becoming a key task for wireless communications. This thesis is devoted to modelling and mitigating interference in wireless networks by dealing with some practical challenges and problems, thereby paving the way to the future.

The rest of this chapter is organised as follows. Section 1.1 will introduce some of the challenges of modern wireless networks that stand behind the motivations for the work included in this thesis. Section 1.2 will explain the thesis layout as well as the accompanying contributions.

1.1 Motivations

One of the objectives of this thesis is to increase the achievable data rates or degrees of freedom (DoF) in wireless communication networks. In the past decades, with better understanding of interference in multi-user wireless networks, achieving higher data rates by increasing the transmit

power became unacceptable. Therefore, a considerable number of studies were conducted to increase data rates in more effective ways. With the development of multiple-input-multiple-output (MIMO)-based techniques, DoF analysis is pushed to the cusp by interference alignment (IA), which has evolved out of [3–6] and was fully introduced in [7]. However, there are two major challenges to implementing IA in practice. On the one hand, obtaining high-quality current channel state information at the transmitter (CSIT) is a difficult task. On the other hand, the overhead of IA increases dramatically with the number of users, i.e., IA is difficult to implement with a large group of users. In contrast, the MAT[1] scheme proposed in [8] is much more practically appealing. It exploits only delayed CSIT to achieve DoF gain for a general number of users. Moreover, the closed-form solutions for the MAT scheme have been proposed for various scenarios. Intuitively, it is meaningful to extend the MAT scheme to more generalised scenarios or to enhance its performance.

Another motivation of this thesis is explained as follows. As the scale of wireless networks is becoming larger, such as MANETs and cellular networks with coverage extensions, the traditional methods seem to be insufficient to evaluate new techniques in such wireless networks for the following three reasons: 1) The signal-to-interference-noise ratio (SINR) should be considered rather than the signal-to-noise ratio (SNR) because interference in these cases is dominant; 2) The effect of path-loss should be considered as it has significant impact on the received signals when the scale of the network is large; 3) Uncertainty in such networks should be taken into consideration in order to adapt to the random data traffic. One effective way to tackle these problems is to develop the stochastic geometry technique. Although there have been a great number of studies in this area, most have focused only on simple settings such as single-input-single-output (SISO), homogeneous distribution, or without a specific interference cancellation technique. Inspired by [9] and [10], evaluating interference cancellation techniques in randomly distributed wireless networks with stochastic geometry can not only help to better understand their performance but also tailor them according to the transmitter and receiver distributions. As IA is a promising beamforming approach to achieve high data rates, it is natural to evaluate its performance and tailor it for randomly distributed large-scale wireless networks.

[1] Named after M. A. Maddah-Ali and D. Tse.

1.2 Thesis Layout and Contributions

Chapter 2 provides a technical background for the work in the thesis. This chapter starts from the components of wireless channels, i.e., the small-scale fading effect and the path-loss effect. Afterwards, fundamentals of MIMO systems are introduced including a toy example of the IA beamforming technique. Finally, locations of the BSs in a central London area are depicted to demonstrate that the BSs in modern large-scale cellular networks are randomly distributed, which leads to the concepts, models and metrics of stochastic geometry.

Chapter 3 enhances the data rate achieved by the MAT scheme for the 2-user and 3-user multiple-input-single-output (MISO) broadcast channel (BC) in the finite SNR case, where the BS has *two* antennas. Firstly, the beamforming vectors and the minimum interference leakage receivers in the 2-user MAT scheme are generalised to endow them with more flexibility. Secondly, by studying the achievable sum rate with the generalised minimum interference leakage receivers, a probabilistic constraints are formalised to guarantee a high probability that the interference leakage power is below a certain threshold. Meanwhile, the objective function is selected to maximise the desired signal detectability. It is worth mentioning that the closed-form solution for the minimum interference leakage receivers is provided. Finally, it is illustrated that each effective MIMO channel for the 3-user case can be decomposed into multiple effective MIMO channels in the 2-user case. That is to say, the proposed probabilistic-constrained optimisation approach can also be utilised to enhance the achievable sum rate for the 3-user case. The simulation results show that the proposed optimisation approach significantly outperforms the MAT scheme especially when the spatial correlation at the receiver side is large.

Chapter 4 extends the MAT scheme to time-correlated channels. Strictly speaking, the major objective of this chapter is to study the DoF regions and the corresponding achievability schemes for the following three scenarios: 1) 2-user interference channel (IC) with 2-antenna BS using imperfect current and delayed CSIT; 2) K-user BC with K-antenna BS using imperfect current and perfect delayed CSIT; 3) 3-user BC with 2-antenna BS using imperfect current and perfect delayed CSIT. Note that the DoF regions are visualised for the 3-user cases for the first time in the literature. Unlike those for the 2-user cases, the convexity of the 3-dimensional DoF regions cannot be straightforwardly observed through the coordinates of vertices. The visualised DoF regions provide an intuitive insight into how the achievable DoF compromise between the three users. It is worth mentioning that all of the DoF regions are consistent with the MAT scheme

and the simple zero-forcing (ZF) beamforming scheme in the presence of perfect current CSIT. In other words, all of the proposed schemes are optimal in the infinite SNR case.

Chapter 5 evaluates the performance of the ZF receiver in multiple-antenna Poisson distributed wireless networks. Note that the performance of analysed receiver represents the average performance of all of the receivers in the system.

- In the first section, the average SINR and transmission capacity achieved with the simple ZF receiver exploiting part of its spatial receive degrees of freedom (SRDoF) to cancel interference in an area are studied. The derived expressions disclose how the SRDoF allocation policy affects the performance in various scenarios.

- In the second section, the impact of the feasibility condition of the ZF receiver to its performance is studied, where the probability that the feasibility condition is satisfied becomes a function of the transmitter distribution. Transmission capacity is derived and is averaged over two instances in which the ZF receiver can or cannot feasibly cancel all the interference in an area.

- In the last section, dynamical data traffic is considered and modelled by queueing theory. For the first time, queueing theory is incorporated with stochastic geometry technique to model both the dynamic data traffic and the random distribution of a wireless network. Because it is unlikely that all of the transmitters and receivers are active in a network, the queueing model can help to more practically characterise the wireless networks by considering the independent joining and leaving of transmitter-receiver pairs. Moreover, the transmitters use the IA approach sequentially to maximise the overlapping of interference signals. With the IA beamforming, the evolving outage probability at the receiver is derived based on the queueing model. Closed-form expression for the transmission capacity is also provided, which provides clear insights into the effectiveness of various system parameters to the transmission capacity, such as the number of data streams and the maximum interferer intensity to maintain a given outage probability.

Chapter 6 considers a generic two-tier heterogeneous wireless network, in which the transmitters in the first tier are distributed as a Poisson point process (PPP) and the transmitters in the second tier are distributed as a Poisson clustered process (PCP). Transmitters and receivers in the second tier are assumed to have imperfect channel state information (CSI) of the nearest cross-tier transmitter. With this imperfect CSI, a distance-dependent IA approach is designed. Specifically, the second-tier transmitters partially align their interferences to the overheard nearest cross-tier

interference at the receivers in the same cluster. The degree of alignment is based on the distance to the nearest cross-tier transmitter, which is considered to be statistically and accurately known. Second-tier receivers use the ZF method to mitigate part of the nearest cross-tier interference as well as the intra-cluster interference aligned to it. Essentially, the distance-dependent IA approach maximises the throughput by compromising between increasing the number of data streams and enhancing the signal-to-interference ratio (SIR) at each data stream. The inter-cluster interference and the marginalised cross-tier interference are modelled by stochastic geometry technique. It is worth mentioning that the near closed-form expression for the outage probability affected by Poisson distributed interferers excluding the nearest one is derived. This expression provides accurate insights into how each factor affects the outage probability if the receiver is served by the nearest transmitter and all the other transmitters are Poisson distributed.

Finally, Chapter 7 provides conclusions based on the work presented in the main chapters and discusses possible future research directions.

Chapter 2
Background

This chapter provides a background of the thesis, starting with the properties of wireless channels. Then, the MIMO system model and its properties are introduced, followed by a brief introduction of the IA beamforming scheme is provided. In the last section of this chapter, the fundamental concepts and models of stochastic geometry are introduced.

2.1 Wireless Channel

Studying the propagation of radio signals in the wireless channels is the primary challenge in wireless communications. It is much more complex than any other communication system because of the physical properties of electromagnetic waves. Roughly, large-scale fading and small-scale fading should be considered in order to characterise the mobile wireless channel.

Most previous work on beamforming schemes focuses only on the small-scale fading [11–14]. It is appropriate when the strength of interference is comparable to the desired signal and the objective is to design an efficient communication system. However, when dealing with relevant issues such as large-scale wireless network planning, consideration of the large-scale fading is of critical importance.

2.1.1 Fading

Fading is defined as the deviation of the attenuation affecting signals over a certain propagation medium. Some related terminologies used in this thesis are given as follows [15]:

- **Large-scale fading:** When the receiver moves through a distance of the order of hundreds of wavelengths, the received signal is affected by *large-scale fading* consisting of *path-loss*, which is a function of distance, and the *shadowing* effect caused by large objects such as buildings and hills.

  - **Path-loss:** This is the mean signal attenuation in terms of the distance between the transmitter and the receiver. The path-loss is the deterministic part of a channel and is well studied in the urban areas in [16]. Path-loss is of critical importance in this thesis and, thus, will be discussed in detail in the next subsection.

- **Small-scale fading:** With the distance between the transmitter and the receiver changing in the order of wavelength, the received signal changes dramatically due to constructive and destructive interference caused by the multi-path effect between the transmitter and the receiver. This effect is usually called *small-scale fading*, or *fading* for simplicity.

  - **Rayleigh fading:** Rayleigh fading is a reasonable model when there are many objects that scatter the signal before it arrives at the receiver. Fundamentally, it is the envelope of a circularly symmetric complex normal distributed random variable. Statistical information of Rayleigh fading can be found in [17].

  - **Nakagami-m fading:** Although Nakagami-m fading is out of the scope of this thesis, it is a very useful model for large-scale wireless networks. When the parameter m increases from one to infinity, the Nakagami-m fading model varies from the Rayleigh fading model to the path-loss only model [18, 19].

2.1.2 Path-Loss Effect

The path-loss is the reduction in power density of an electromagnetic wave as it propagates through space. According to the laws of physics in free space, the received power will decay as l_{ij}^{-2} where l_{ij} is the distance between Tx-j and Rx-i. However, because the reflected signals reduce the direct signal and the obstacles might also absorb some power, the received power will decay considerably faster than l_{ij}^{-2}. According to empirical evidence, the received power decays as l_{ij}^{-2} near the transmitter and decays exponentially at a distance far from the transmitter.

The inverse exponential relationship between received signal power and distance is commonly known, i.e., $\frac{P_r}{P_t} \propto l_{ij}^{-\varsigma}$, where ς is called the *path-loss exponent* and P_r, P_t are the received and transmit power with units in Watts (W), respectively. Although a singular point exists in this path-loss model, its tractability has attracted a considerable amount of work, such as [9, 10, 20–23]. Table 2.1 lists the empirical path-loss exponent values in various environments.

Environment	Path-Loss Exponent α
Infinite Free Space	2.0
Urban Area	2.7 to 3.5
Suburban Area	3.0 to 5.0
Office	2.6 to 3.0
Store	1.8 to 2.2

Table 2.1: Path-loss exponent values in various environments [1].

2.2 MIMO System

The initial idea of MIMO dates back to the work of Kaye and George (1970), Branderburg and Wyner (1974) [24] and van Etten (1975, 1976). From 1984 to 1986, Bell Laboratories published their work on beamforming techniques [25], which is known as the prototype of the modern beamforming techniques.

Generally speaking, MIMO system represents a communication system whereby both transmitters and receivers are equipped with multiple antennas to improve the overall performance. MIMO has attracted attention because it can increase the capacity and transmission range without increasing the transmit power or bandwidth. The transmitter encodes the symbols and emits them with appropriate weights on each of the antennas, which is known as *beamforming*. The signal propagates through the channel, where fading effects dominate, and arrives at the receiver. The receiver, on the other side, retrieves the signal from the antennas and decodes the desired symbols.

Straightforwardly, the MIMO system of a transmitter-receiver pair, Rx-i and Tx-i, is illustrated in Figure 2.1. In this figure, the MIMO system has M transmitter antennas, N receiver antennas and $\bar{d}$ data streams. The post-processing signal Rx-i can be modelled by the following equation:

$$\hat{\mathbf{y}}_i = \mathbf{U}_i^H \mathbf{H}_{ii} \mathbf{V}_i \mathbf{s}_i + \mathbf{U}_i^H \mathbf{n}_i, \tag{2.1}$$

where $\hat{\mathbf{y}}_i \in \mathbb{C}^{\bar{d}\times 1}$ is the post-processing received symbol vector at Rx-i and $\mathbf{s}_i \in \mathbb{C}^{\bar{d}\times 1}$ contains the symbols transmitted by Tx-i. The channel matrix $\mathbf{H}_{ii} \in \mathbb{C}^{N\times M}$ contains entries representing channel gains from each of the transmitter antennas to each of the receiver antennas. The beamforming matrix and the receiving matrix are $\mathbf{V}_i \in \mathbb{C}^{M\times \bar{d}}$ and $\mathbf{U}_i \in \mathbb{C}^{N\times \bar{d}}$ respectively. Symbol $\mathbf{n}_i \in \mathbb{C}^{N\times 1}$ stands for additive white noise vector.

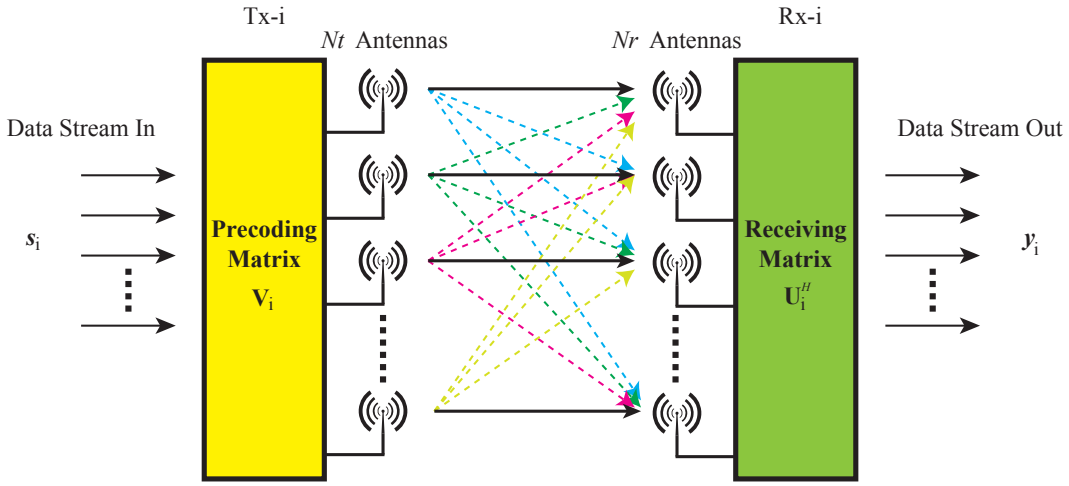


Figure 2.1: A point-to-point MIMO wireless communication system.

2.2.1 CSIT and Beamforming Matrix Acquisition

In modern wireless networks, CSIT acquisition is carried out mainly through training and feedback processes. According to [26], the CSIT acquisition can be generally accomplished through the following steps:

1. Common training: The transmitter sends shared training symbols by which the receiver can estimate the channel using its observations of the symbols.

2. CSI feedback: The receiver feeds back its estimated CSI to the transmitter after completing the common training.

3. Beamformer selection: The transmitter will design the beamforming matrix according to the obtained CSI.

4. Dedicated training: The transmitter will transmit new shared symbols with the designed beamforming matrix. Since the receivers have the CSI already, this step is dedicated to sharing the beamforming matrix with the receiver.

However, in most cases, instantaneous and perfect CSIT is not available for two reasons:

- **Feedback Error:** There is only limited bandwidth in the CSI feedback link between the receiver and the transmitter, which restricts the accuracy of the CSIT.

- **Delay:** There is a delay between the CSI measurement at the receiver and the CSI utilisation at the transmitter. In a time-varying channel, the delay causes the available CSI to be

different from the current channel. If the coherence time of the channel is not too short compared to the symbol period, there is a correlation between the available CSI and the current channel. This time-correlation of channel will be exploited in Chapter 3.

2.2.2 Broadcast Channel and Interference Channel

Generally speaking, there are several kinds of multiuser communication model, such as IC, BC, interference broadcast channel (IBC), X channel. The models of IC and BC are shown in Figure 2.2, in which the black solid lines represent the desired signals and the dotted lines stand for interference signals. Note that, for BC, interference is also overheard at the receivers. In practice, BC models the situation where a single transmitter sends independent information to different receivers such as cellular networks. In contrast, IC models the situation where each of the transmitters sends independent information to one intended receiver.

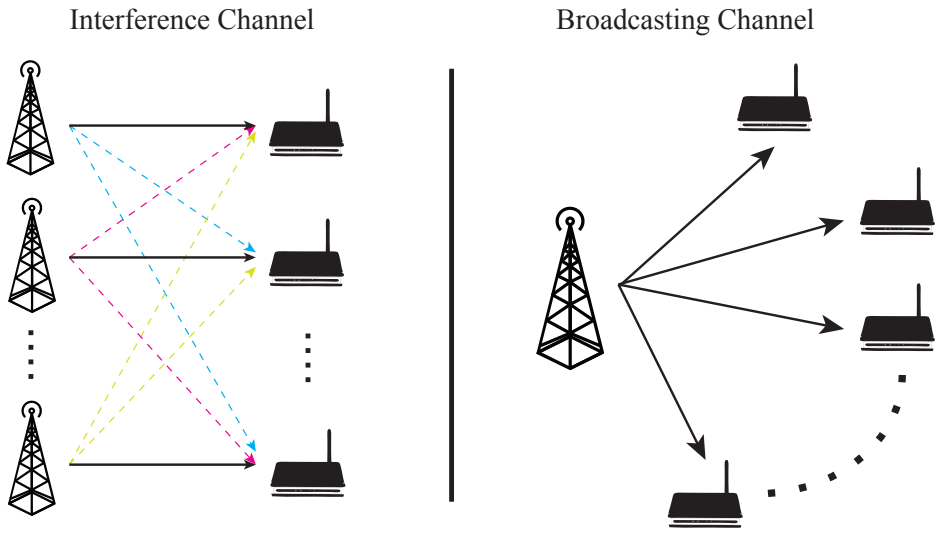


Figure 2.2: Broadcast channel and interference channel.

There are two explicit differences between IC and BC for in a K-user case:

- There are K^2 potential links in the IC in total but only K in the BC, i.e., channel estimation is more complex in the IC.
- In the BC, all symbols are transmitted by the BS, i.e., it can reconstruct the interference overheard at the receivers in the presence of delayed CSIT. This property of the BC helps to design schemes in Chapter 4.

2.2.3 Interference Cancellation

In wireless communications, *interference* typically refers to the addition of unwanted signals to the desired signal. Theoretically, a receiver overhears interference if there are undesired transmitters working at the same frequency as the desired transmitter. Although interference can be easily avoided by allocating frequency bands to the transmitters, this consumes large amounts of spectrum resources. As this thesis focuses on enhancing the spectrum efficiency, all transmitters are assumed to work at the same frequency.

In a multiuser communication system, maximising the data rate for each receiver independently is hardly the optimal choice to achieve the maximum sum rate. This phenomenon can be observed in Figure 2.3 which plots the achievable sum rates for the 2-user MISO IC with 2-antenna transmitters versus SNR. Three different schemes are considered: the ZF beamforming scheme, the Time-Division-Multiple-Access (TDMA) scheme and the greedy beamforming scheme where transmitters design their beamforming matrices in order to maximise their own data rates. Rayleigh fading is considered and each point is averaged over $1,000$ channel realisations. In this simulation, variance of the CSI estimation error is defined as $P^{-\alpha}$, where P is the SNR. The condition $\alpha = 0.3$ is considered for the low CSIT quality ZF line and $\alpha = 0.6$ for the high CSIT quality ZF line as well as the greedy beamforming line. The most important observation from this figure is that the line of greedy beamforming scheme is the only one that does not continuously increase with the SNR. The reason behind this is that the interference power increases at the same rate as the desired signal power. This observation illustrates that it is meaningless to infinitely increase the transmit power, and more effective ways to increase the spectrum efficiency are expected.

For a MIMO system, there are several receiver structures to mitigate or cancel the interference signals, e.g., minimum mean square error (MMSE) receiver, successive interference cancellation (SIC) receiver and ZF receiver. The ZF receiver (also known as the decorrelator), which is considered in Chapter 5 and Chapter 6, is a projection operation followed by a matched filter. Specifically, if Tx-j is causing interference to Rx-i, where $j \neq i$, and Rx-i has perfect current CSI between them, i.e., $\mathbf{H}_{ij}$, the ZF receiver Rx-i can design its receiving matrix so that $\mathbf{U}_i^H \mathbf{H}_{ij} \mathbf{V}_j = \mathbf{0}$. If $\mathbf{P}$ is the basis of the null space of cancelled interference at the q-th data stream of Rx-i (including interference from transmitters other than Tx-i and other data streams from Tx-i), the corresponding decorrelator of the q-th data stream, denoted as $\mathbf{f}_q$, is given by $\mathbf{f}_q^H = \frac{\bar{\mathbf{h}}_q^H \mathbf{P}\mathbf{P}^H}{|\bar{\mathbf{h}}_q^H \mathbf{P}\mathbf{P}^H|}$, where $\bar{\mathbf{h}}_q$ is the q-th column of $\mathbf{H}_{ii}\mathbf{V}_i$. It is worth mentioning that the quality of

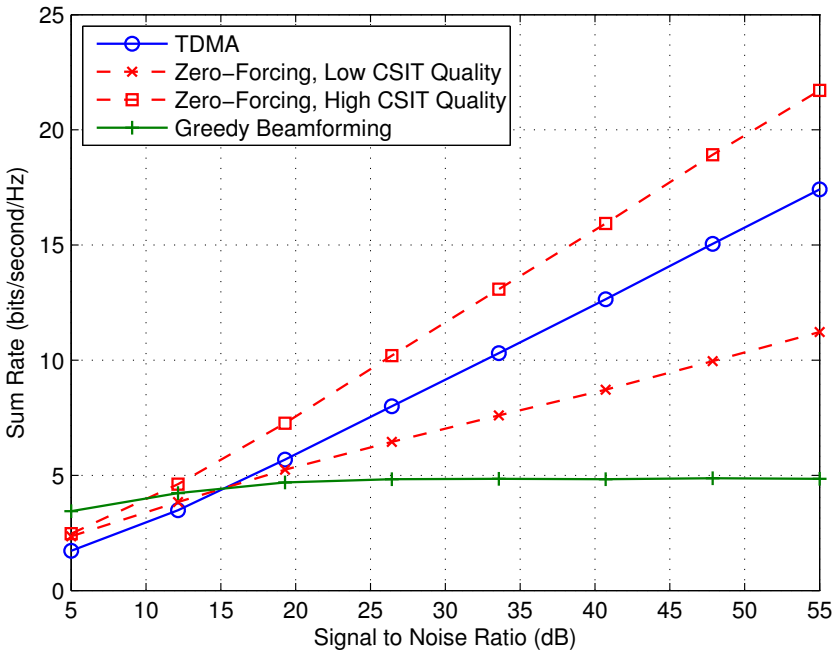


Figure 2.3: Sum rates of the ZF scheme, the TDMA scheme and the greedy beamforming scheme.

current CSI plays an important role to the ZF receiver.

2.2.4 Degrees of Freedom Region

In a communication system, DoF is a popular information-theoretic performance metric indicating the number of interference-free data streams that can be communicated over the wireless channel. To be more specific, for a single transmitter-receiver pair, the achievable DoF is defined as the pre-log factor of the capacity when the SNR tends to be infinite. If P is the SNR, the DoF can be expressed as,

$$\lim_{P\to\infty} \frac{C(P)}{\log P}, \tag{2.2}$$

where $C(P)$ is the channel capacity with SNR P.

The importance of DoF study is that the channel capacity increases almost linearly with DoF in the high SNR range. Defining d_i as the DoF achieved by Rx-i, the DoF tuple in a K-user wireless network can be expressed as,

$$\begin{aligned}&\Big\{(d_1,d_2,\ldots,d_K)\in\mathbb{R}_+^K|\forall(x_1,x_2,\ldots,x_K)\in\mathbb{R}_+^K, x_1d_1+x_2d_2+\cdots+x_Kd_K\\ &\leq \limsup_{P\to\infty}\left[\sup_{\mathbf{R}(P)\in\bar{C}(P)}[x_1R_1(P)+x_2R_2(P)+\cdots+x_KR_K(P)]\frac{1}{\log(P)}\right]\Big\},\end{aligned} \tag{2.3}$$

where $R_i(\cdot)$ is the data rate at the i-th receiver, $\mathbf{R}(\cdot)$ is the rate tuple and $\bar{C}(\cdot)$ is the capacity region. DoF is the core research object in Chapter 4.

DoF Region of 2-User MISO Broadcast Channel

The DoF region of the 2-user MISO BC with M-antenna BS ($M \geq 2$) is illustrated in Figure 2.4 in the absence of any delayed CSIT. From this figure, it can seen that with perfect current CSIT, each receiver obtains *one* DoF. That is to say, the BS can simultaneously send one data stream to each receiver with interference power submerged in white noise. When the quality of the current CSIT decreases, interference cannot be fully cancelled, which leads to DoF degradation. When the quality of the current CSIT decreases continuously, the TDMA scheme can help to guarantee *one* DoF in total, which is represented by the line between points $(0, 1)$ and $(1, 0)$.

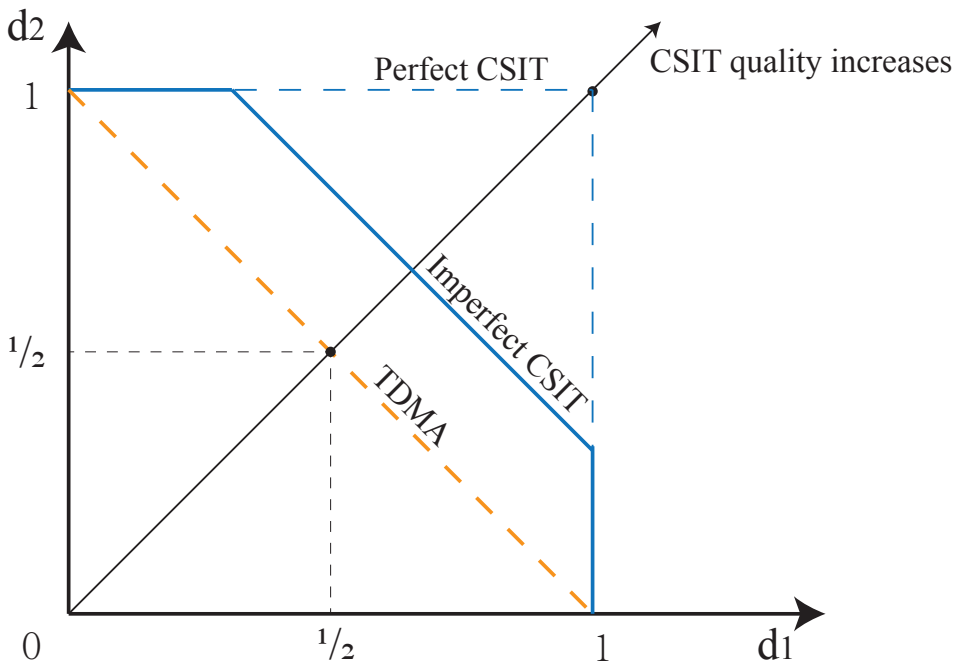


Figure 2.4: Degrees-of-freedom region of a 2-user MISO BC.

2.2.5 A Brief Introduction to Interference Alignment

Interference alignment refers to the overlapping of multiple interference signals at interference subspaces of receivers so that the interference-free dimension of the desired signal is maximised. In time-varying IC, [7] uses IA over time extension to achieve $\frac{1}{2}$ DoF at each receiver, which is half of the total achievable DoF. This is quite an exciting result because the total DoF are linearly

increasing with the number of users, i.e., the total DoF achieved in the K-user IC is $\frac{K}{2}$. With the aforementioned property, IA has been widely considered in literature, e.g., in relay channel [27], in cellular network [28], in Orthogonal-Frequency-Division-Multiplexing (OFDM) system [29], with diversity [30], and with a spatial correlation [31]. As closed-form solutions of IA beamformers are rarely available, it is usually implemented using a retrospective algorithm [32]. There are some other iterative algorithms used to implement IA, such as the max-SINR approach [33] and the minimum interference leakage approach [34].

IA has three major variants: IA over a time-extension channel, IA over a frequency-extension channel and IA over a MIMO channel. IA over a time-extension channel is merely a theoretical tool to study the property of IA because non-causal knowledge of CSI is required. IA over a frequency-extension channel is applicable but requires global synchronisation of frequency, which is a difficult task in modern wireless networks. IA over a spatial dimension channel is the most attractive variant because only current CSI over a MIMO channel is required. Therefore, this thesis focuses only on IA a MIMO channel, namely *MIMO IA*.

Challenges of Interference Alignment

As a promising technique to achieve higher DoF, IA faces two major challenges:

- As stated in Section 2.2.1, perfect and instantaneous current CSIT is not available based on the modern CSIT acquisition method.
- For the IA over a time-extension channel, it has been proven that $\frac{K}{2}$ DoF can be achieved in total for the K-user case. However, for MIMO IA, the limited number of antennas restricts the effectiveness of IA. Actually, the feasibility condition of IA [35] states that the following inequality must hold in order to implement MIMO IA feasibly in a K-user IC:

$$N + M - \bar{d}(K + 1) \geq 0. \tag{2.4}$$

  Recall that M is the number of transmitter antennas, N is the number of receiver antennas and $\bar{d}$ is the number of data streams transmitted between each transmitter-receiver pair.

A Toy Example of MIMO Interference Alignment

Consider a 3-user IC in which all of the transmitters and receivers are equipped with 3 antennas. In order to visualise the directions of signals, all of the symbols, entries of the beamforming vectors and the channel matrices are assumed to be real numbers. All of the transmitters and receivers have global perfect current CSI. Tx-1 transmits *two* encoded data streams along two beamforming vectors, as shown in the top left sub-figure of Figure 2.5. Tx-2 and Tx-3 send their independently encoded data streams along their beamforming vectors, as shown in the two bottom left sub-figures. The directions of the beamforming vectors are designed so that after multiplying the channel matrices, the directions of the received signals are shown on the right-hand side of Figure 2.5. Specifically, interference signals from Tx-2 and Tx-3 are aligned at Rx-1. The interference signal from Tx-2 is aligned with one of the two data streams from Tx-1 at Rx-3. Similarity, the interference signal from Tx-3 is aligned with one of the two data streams from Tx-1 at Rx-2. Therefore, each receiver is able to decode its own symbol(s) without interference by projecting all of the received signals to the null space of its interference subspace. As a result, Rx-1 achieves *two* DoF, and Rx-2 and Rx-3 achieve *one* DoF. In total, the system achieves *four* DoF in *one* time slot.

2.3 Stochastic Geometry

Stochastic geometry technique deals with random spatial patterns, where point processes are the most important and typical objects. According to [36–38], stochastic geometry technique is devoted to the study of connectivity and signal power in wireless networks where the transmitters and receivers are placed according to certain distribution. Note that the connectivity study is out of the scope of this thesis.

2.3.1 Wireless Networks and Point Process

Point processes in this thesis are random collections of points that reside in 2-dimensional spaces. Each point represents the location of one node (transmitter or receiver) in wireless networks in the real world. Figure 2.6[1] demonstrates the locations of the BSs in a central London area.

Clearly, the locations of BSs in large cities such as the 52 BSs in a central London area do not

[1]Resource website: http://sitefinder.ofcom.org.uk/search

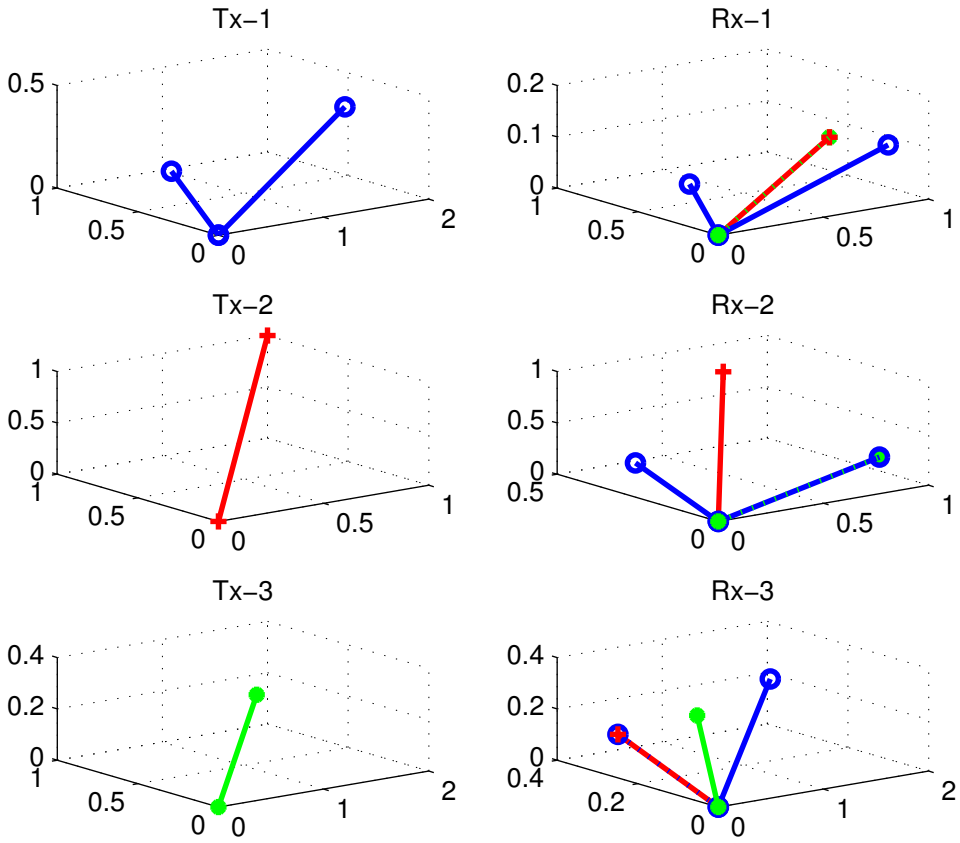


Figure 2.5: An example of MIMO IA for a 3-user IC.

form a deterministic pattern or lattice. Therefore, a random approach such as stochastic geometry to model the BS distribution is sensible.

2.3.2 Borel σ-algebra

A 2-dimensional Borel set $\mathcal{B}^2$ is any set in a 2-dimensional topological space that can be formed by taking the complement, countable unions and intersections of closed, open or half-open sets. A collection of all Borel sets forms the *Borel σ-algebra*.

If a set consists of a countable number of singletons, it is a countable set. In a 2-dimensional space, a set $A \subseteq \mathbb{R}^2$ is said to be bounded if there exists a point $x \in \mathbb{R}^2$ and value $r \in \mathbb{R}$ so that a ball centred at x with radius r, denoted as $b(x, r)$, has $A \subset b(x, r)$. If a bounded set is also closed, it is *compact*.

Figure 2.6: Locations of BSs near River Thames, central London.

2.3.3 Measurability in 2-Dimensional Space

Generally, an ordered pair $(A, \mathcal{A})$ is called a measurable space if $\mathcal{A}$ is a σ-algebra on the subsets of A, where σ-algebra on a set is a collection of its subsets that is closed under countably many set operations such as complement, union and intersection. Note that the point patterns in this thesis, i.e., $(\mathbb{R}^2, \mathcal{B}^2)$, are measurable. A function $f : \mathbb{R}^2 \mapsto \mathbb{R}$ is $\mathcal{B}^2$-measurable if and only if the pre-image of 1-dimensional Borel set $A \in \mathcal{B}$ is an element of $\mathcal{B}^2$, i.e.,

$$f^{-1}(A) = \{x \in \mathbb{R}^2 : f(x) \in A\} \in \mathcal{B}^2.$$

Consider all pairwise disjoint sets $A_i \in \mathcal{A}, i = 1, 2, \ldots, n$ and $n \in \mathbb{N} \cup \{\infty\}$. A measure ν is a function $\nu : \mathcal{A} \to \mathbb{R} \cup \{\infty\}$ with countable additivity, i.e.,

$$\nu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \nu(A_i),$$

and satisfying $\nu(\emptyset) = 0$, where $\emptyset$ is the empty set.

Lebesgue Measure

The Lebesgue measure, denoted as $|\cdot|$, is a commonly used measure in this thesis and is basically a measurement of volume in the Euclidean space. For example, the Lebesgue measure of an area of a disk with radius 3 centred at point x in a 2-dimensional Euclidean space is $|b(x, 3)| = 9\pi$.

Dirac Measure

The Dirac measure, denoted as $\delta_x(A)$, is a measure at point x for a set $A \subset \mathcal{B}^2$ which is equal to 1 if $x \in A$ and 0 otherwise. The Dirac measure can be expressed by indicator function $\mathbf{1}_A(\cdot)$ as:

$$\delta_x(A) = \mathbf{1}_A(x).$$

Counting Measure

For a point process Φ, the counting measure $\Phi(A)$, where $A \subset \mathcal{B}^2$, is to count the number of singletons in Φ that fall in the set A. The counting measure can be expressed as:

$$\Phi(A) = \sum_{x \in \Phi} \delta_x(A).$$

It will be shown that the counting measure is the most important measure to describe a 2-dimensional point process.

Intensity Measure

The intensity measure $\Lambda(\cdot)$ for a set A, where $A \subset \mathcal{B}^2$, can be expressed by the following equation:

$$\Lambda(A) = \mathbb{E}[\Phi(A)],$$

where $\mathbb{E}[\cdot]$ is the expectation operation. The intensity measure can be explained as the expected number of points falling in a set.

Campbell's Theorem

Campbells theorem refers to an equation relating to the expectation of a function summed over a point process to an integral involving the intensity measure of the point process. For a function $f : \mathbb{R}^2 \mapsto [0, \infty)$ which is a measurable function and Φ is a point process, according to [39], Campbell's theorem can be expressed as:

$$\mathbb{E}\left[\sum_{x \in \Phi} f(x)\right] = \int_{\mathbb{R}^2} f(x)\Lambda(\mathrm{d}x).$$

A specific proof of the Campbell's theorem can be found in Theorem A.2 of [39].

2.3.4 Fundamentals of Point Process

Definition

According to [37], a point process Φ is a countable random collection of points that reside in some measurable space, usually the 2-dimensional Euclidean space. The associated σ-algebra consists of the Borel sets $\mathcal{B}^2$, and the measure used is the Lebesgue measure.

Formalism

There are two main ways to formalise a 2-dimensional point process: the *random set formalism* and the *random measure formalism*. In a random set formalism, a point process is regarded as a countable set consisting of random variables. In this thesis, the random measure formalism is considered, which characterises a point process by the counting measures of sets.

2.3.5 Poisson Point Process

A PPP assumes total independence of the locations of nodes. That is to say, the location of any node is completely uncorrelated with the locations of all other nodes. If a point process Φ is a homogeneous PPP, i.e., a Poisson process with the same intensity λ all over the Euclidean plane, then Φ must have the following properties:

- For any compact set A, $\Phi(A)$ has Poisson distribution with mean value $\lambda|A|$;
- If $A_1, A_2, \ldots, A_n$ are disjoint bounded sets, then $\Phi(A_1), \Phi(A_2), \ldots, \Phi(A_n)$ are independent random variables.

In a 2-dimensional space, PPP can be characterised by counting measure as the following equation:

$$\Pr\big(\Phi(A) = k\big) = \frac{\exp\big(-\lambda|A|\big)\big(\lambda|A|\big)^k}{k!},\ k \in \mathbb{N},$$

and its pattern can be seen from Figure 2.7. Note that $\Pr(\cdot)$ means the probability hat an event happens. PPP is the most fundamental point process and will be studied in Chapter 5.

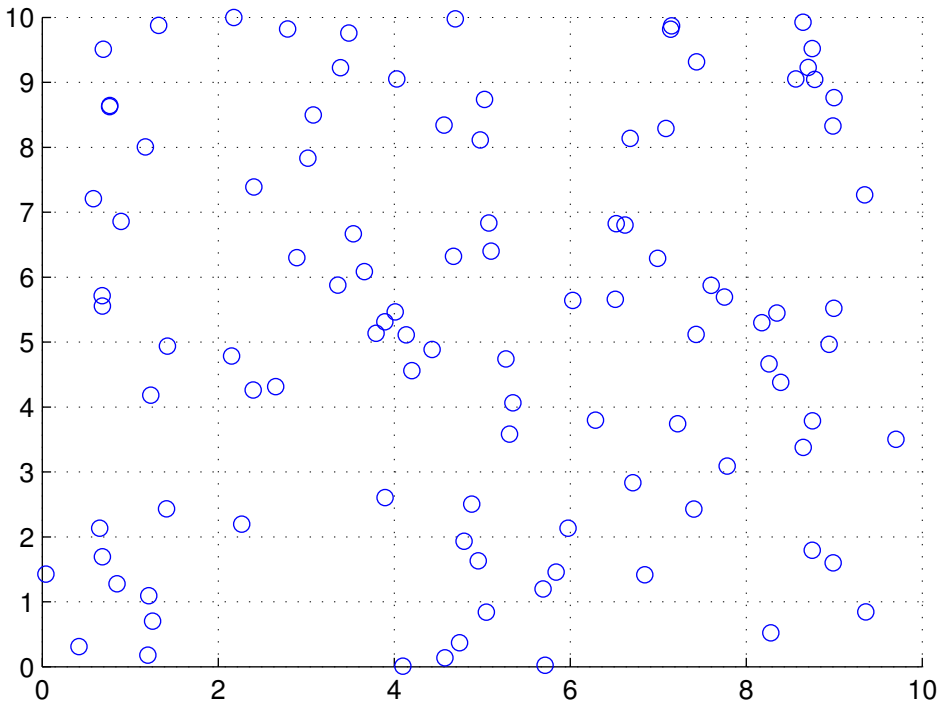


Figure 2.7: A poisson point process with unit intensity.

2.3.6 Poisson Clustered Process

A clustered process assumes that the location of a node is correlated with some other nodes. Usually, a clustered point process is generated by a parent point process and daughter point processes. The points in the parent process are typically not included in the clustered process but work as reference points (centres) for the daughter processes. The parent point process is denoted as $\Phi_P = \{x_0, x_1, x_2, \dots\}$ where point $x_i, \forall i \in \mathbb{N}$ is the reference point of cluster with index i. Denoting $\bar{\phi}_i, i \in \mathbb{N}$ as the untranslated daughter process (with parent point at the origin) of cluster i, the translated daughter process (centre shifted to the corresponding parent point) can be denoted as $\phi_i = \bar{\phi}_i + x_i$. Hence, if a point process Φ is a clustered process with parent process Φ_P and untranslated daughter processes $\bar{\phi}_i, \forall i \in \mathbb{N}$, this point process can be expressed as:

$$\Phi = \bigcup_i (\bar{\phi}_i + x_i) = \bigcup_i \phi_i,$$

If the parent process is stationary and the daughter processes are independent and identically distributed (i.i.d.), the clustered process is called the *homogeneous independent clustered process*. The *Neyman-Scott clustered process* is a homogeneous clustered process wherein the parent

process is a PPP and the daughter points in each cluster are random in number, independent of each other, and identically distributed. The Neyman-Scott clustered process is generally called PCP. Furthermore, the *Matérn process* is a Neyman-Scott process wherein nodes in each daughter process are uniformly scattered in a ball centred at the parent point. A Matérn process is shown in Figure 2.8 with 0.01 parent point intensity, 3 daughter points and 5 cluster radius. The *Thomas process* is a Neyman-Scott process wherein nodes in each daughter process are Gaussian distributed centred at the parent point.

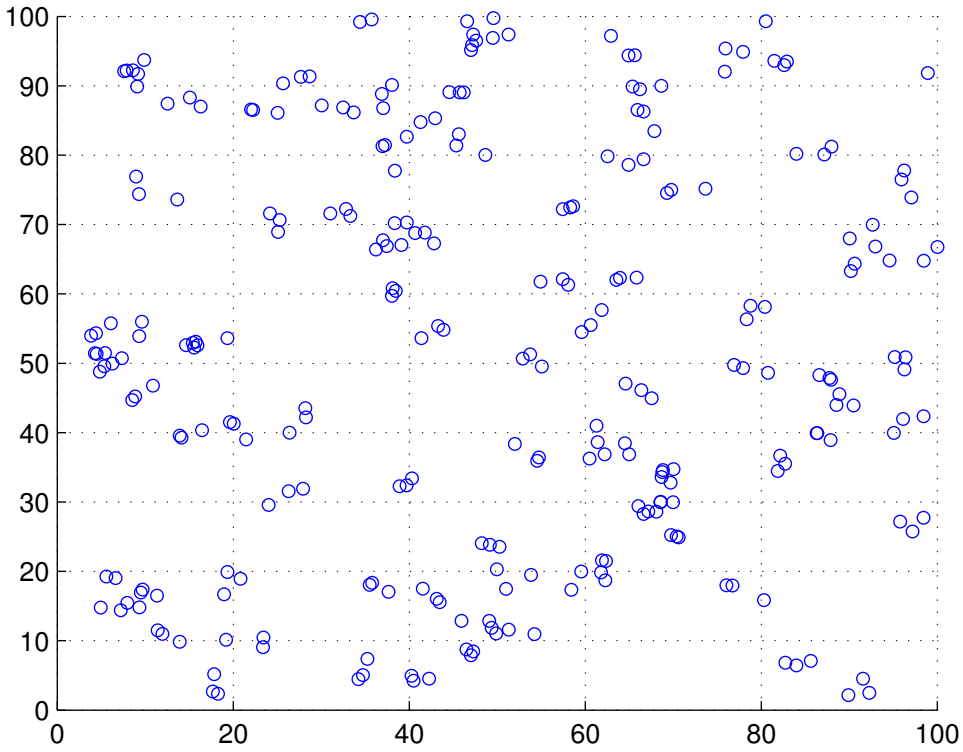


Figure 2.8: A Matérn process in the Euclidean plane.

This thesis considers a Matérn process in each cluster in Chapter 6.

2.3.7 Poisson Clustered Process and Heterogeneous Wireless Networks

In a modern LTE-A cellular network, various kinds of wireless infrastructure are deployed to increase the network capacity [40] including the macro BSs, the pico BSs, the femto BSs and the relays. Moreover, this heterogeneity is widely believed to continue in the next generation of wireless networks as discussed in [41]. For an LTE-A network, macro BSs can be reasonably modelled by PPP [42]. The distribution of pico BSs and terminals are more like PCP because more

services are required in crowded places such as parks, shopping malls and stations. An incomplete distribution of devices in a heterogeneous cellular network is shown in Figure 2.9.

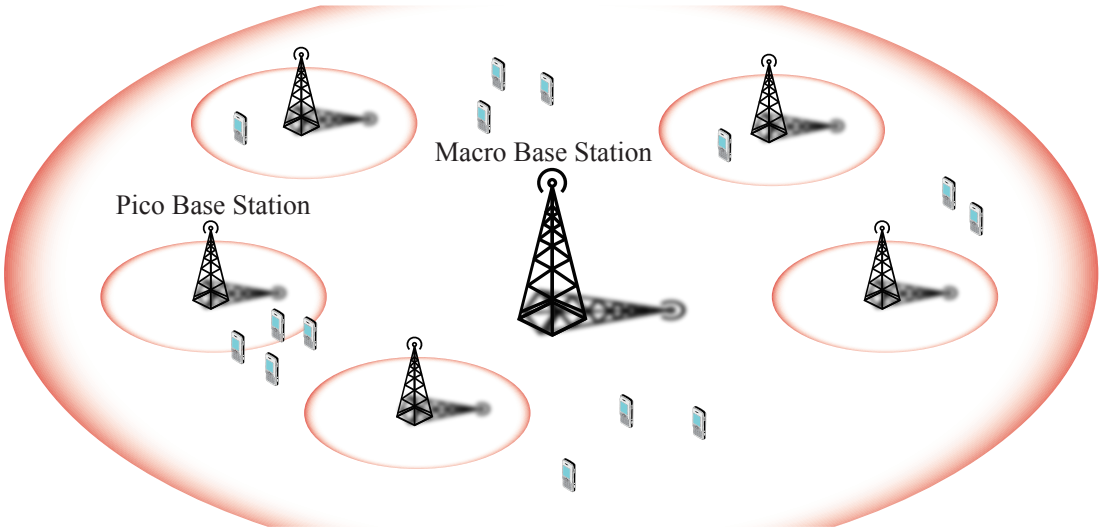


Figure 2.9: The BSs and terminals in a heterogeneous cellular network.

2.3.8 Voronoi Cell

The *Voronoi cell* of a point x in a point process Φ, denoted as $V(x)$, consists of locations in the Euclidean space of which the distances to x are not greater than the distance to any other point in Φ, i.e.,

$$V(x) = \{y \in \mathbb{R}^2 : \|x - y\| \leq \|z - y\|, \forall z \in \Phi \backslash \{x\}\}.$$

The *Voronoi tessellation* of Φ divides the space into a number of regions, each of which is the Voronoi cell of a point in Φ. Figure 2.10 illustrates the Voronoi tessellation of a 2-dimensional PPP with unit intensity in a 10×10 area.

The Voronoi cell is a very important tool in wireless network planning. Inside the cell of the i-th BS, the mean signal power from the i-th BS is larger than that of any other BSs due to the path-loss effect. Therefore, BSs typically only serve the terminals inside their own Voronoi cells. If the BSs have different transmit power, a weighted Voronoi diagram should be considered instead, as in [43, 44].

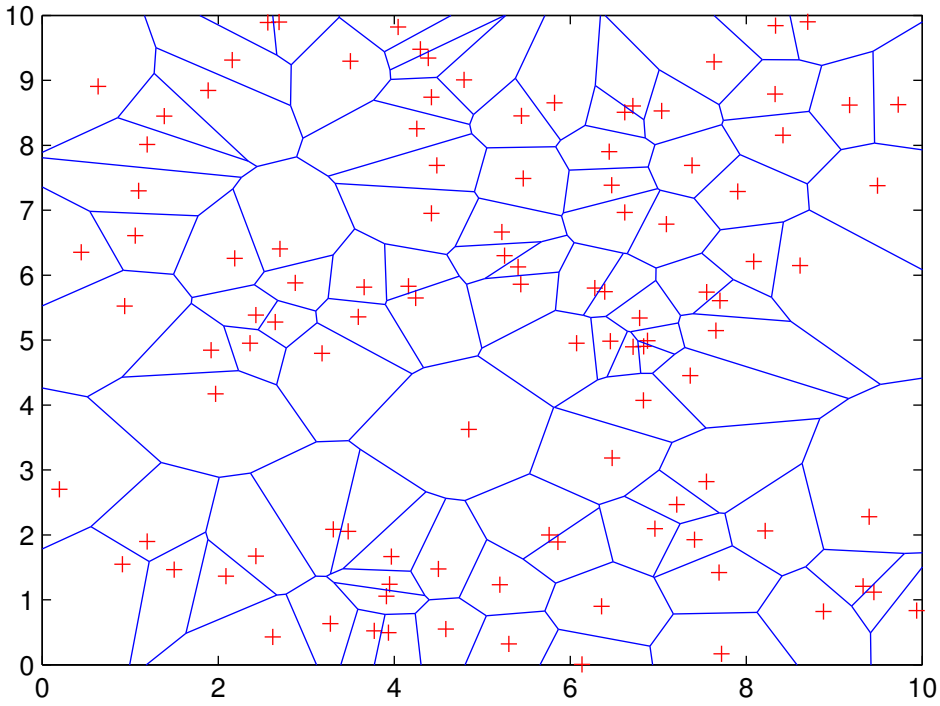


Figure 2.10: Voronoi tessellation of a PPP in 2-dimensional space.

2.3.9 Outage Probability

Generally, a receiver experiences outage when the received signal strength does not support the required data rate due to the channel variations. In the case of slow fading, outage probability, denoted as ϵ, is defined as the percentage of time outage occurs. In this thesis, because outage probability is only analysed in distributed wireless networks where interference is dominant and white noise is negligible, outage probability is formally defined as the probability that the SIR is below a certain threshold, i.e.,

$$\epsilon = \Pr(SIR < \theta),$$

where θ is called the *pre-specified SIR threshold*.

2.3.10 Throughput

Throughput is defined as the rate of successful message delivery over a communication channel. For a data stream with a success probability (one minus the outage probability) constrained as $1-\epsilon$ at data rate R bits/sec/Hz, the throughput is given by $(1-\epsilon)R$. If there are, in total, $\bar{d}$ mutually

independent data streams in a link and all of the data streams have the same success probability and data rate, the average throughput for this link is $\bar{d}(1-\epsilon)R$. As discussed in [9], if there are K independent users in the system and all achieve the same throughput, the total throughput $\mathcal{T}$ of the system can be written as,

$$\mathcal{T} = KR\bar{d}(1-\epsilon), \tag{2.5}$$

where the data rate R is a function of θ.

2.3.11 Transmission Capacity

Transmission capacity is a performance metric tailor-made for randomly distributed wireless networks because it takes the intensity of transmitter-receiver pairs into consideration. Transmission capacity $\mathcal{C}$ is defined as the product of the success probability $1-\epsilon$, the corresponding maximum intensity of transmitters to guarantee the outage probability λ^{ϵ}, the number of data streams $\bar{d}$ and the rate of each data stream R, i.e.,

$$\mathcal{C} = (1-\epsilon)\lambda^{\epsilon}\bar{d}R. \tag{2.6}$$

Note that transmission capacity is closely related to throughput, which is discussed in Section 2.3.10. In fact, throughput characterises the spectrum efficiency of a system with a certain number of users, while transmission capacity is a metric of the spectrum efficiency per unit area in a distributed wireless network, or *area spectrum efficiency*.

Chapter 3

Probabilistic-Constrained Optimisation for MISO BC Using Delayed CSIT

3.1 Introduction

Multiplexing gain achieved by interference management techniques such as multiuser MIMO and IA are heavily based on the CSIT. As stated in Section 2.2.1, CSIT is commonly obtained by feedback from the receivers. Usually, a scheme assuming perfect current CSIT implies that the fed back CSI can be used to perfectly predict the current CSIT. However, the prediction will not be able to achieve any multiplexing gain when the channel coherence time is small compared to the CSI feedback delays. The MAT scheme proposed in [8] exploits only the delayed CSIT and successfully achieves DoF higher than the single DoF in the absence of any CSIT. In other words, the delayed CSIT is still very useful to achieve higher DoF. Note that the MAT scheme actually provides specific achievability schemes in BC for three different system settings: 1) The 2-user with 2-antenna BS setting, denoted as $\mathcal{S}_{MAT}^{(1)}$; 2) The K-user with K-antenna BS setting, where $K \geq 3$, denoted as $\mathcal{S}_{MAT}^{(2)}$; 3) The 3-user with 2-antenna BS setting, denoted as $\mathcal{S}_{MAT}^{(3)}$. Note that all of the aforementioned symbols are for BC unless stated otherwise in this thesis.

Recent studies have been carried out to introduce optimisation tools to the MAT scheme so that higher data rates can be achieved in the finite SNR case. The motivation is that the MAT scheme is too conservative in this scenario as all the DoF are exploited to align the interference [45]. Specifically, in [46] and [45], the beamforming vectors in the MAT scheme are optimised in $\mathcal{S}_{MAT}^{(1)}$ and $\mathcal{S}_{MAT}^{(2)}$ using dual SINR approach, which is realised by efficiently striking a balance between IA and desired signal detectability enhancement. Note that the dual SINR approach is an iterative-free, closed-form and suboptimal solution. The achievable sum rate in [45] is derived by permuting rows in the channel matrices. In other words, the dual SINR scheme is designed in the absence of a specific receiver structure. As the MAT scheme has already provided the specific receiver structure, it is more practically appealing to take it into consideration when designing optimisation

schemes for the MAT scheme. Meanwhile, the dual SINR scheme ignores the unknown channel entries and thereby, causes performance degradation if the receiver structure is considered as stated in [45]. Hence, the primary objective of our proposed optimisation approach is to tackle these two issues. Specifically, when designing the optimisation scheme for the MAT scheme in this chapter, the receiver structure is considered and the statistical information of the unknown channel entries are exploited.

Usually, in order to take the unknown channel entries into consideration, robust optimisation techniques should be considered. Generally, robust optimisation techniques can be categorised into two approaches: the worst-case approach [47–50] and the probabilistic-constrained approach [51,52]. The authors in [51] proposed a probabilistic-constrained approach for the MIMO system, where it was proven that the Gaussian CSI mismatch enabled the conversion of the optimisation problem to a tractable deterministic form with lower computational complexity. After that, some other work also showed that high performance could be achieved using the probabilistic-constrained approach for various settings. For example, in [52], the authors studied a probabilistic-constrained optimisation approach for a narrowband beamformer, and in [53], the probabilistic-constrained optimization approach was studied for a multi-antenna downlink beamformer. As the effective MIMO channels for the MAT scheme are subject to Gaussian channel mismatch in the case of Rayleigh fading, it is interesting to introduce the probabilistic-constrained optimisation approach to the MAT scheme and understand the influence of the probabilistic constraint. This is the secondary motivation of this chapter.

In this chapter, a probabilistic-constrained optimisation approach for $\mathcal{S}_{MAT}^{(1)}$ and $\mathcal{S}_{MAT}^{(3)}$ using perfect delayed CSI is designed. The proposed optimisation approach begins with generalising the MAT scheme for $\mathcal{S}_{MAT}^{(1)}$. Afterwards, the expression of the achievable sum rate is derived to help understand the role of the generalised beamforming vectors in the interference leakage power and the desired signal detectability. It is worth mentioning that the generalised minimum interference leakage receiver and the unknown channel entries are included in the expression. Afterwards, a probabilistic-constrained optimisation problem is formalised, which keeps a high probability that the interference leakage power at the receivers is below a certain threshold while maximising the desired signal detectability. Note that the closed-form solution for the minimum interference leakage receivers is provided. Finally, the proposed optimisation approach is migrated to the MAT scheme in $\mathcal{S}_{MAT}^{(3)}$. Based on the simulation results, the relationship between the allowed interference leakage power and the achievable sum rate is obtained. Meanwhile, the proposed scheme outperforms the MAT scheme and the dual SINR scheme remarkably when the spatial

correlation at the receiver side is large.

The remainder of this chapter is organised as follows. Section 3.2 presents the system model and a brief introduction of the MAT scheme. Section 3.3 proposes the probabilistic-constrained optimisation approach in $\mathcal{S}_{MAT}^{(1)}$. Section 3.4 extends the probabilistic-constrained optimisation approach to $\mathcal{S}_{MAT}^{(3)}$. Simulation results are presented in Section 3.5 and are followed by summaries in Section 3.6.

3.2 System Model

This chapter considers a K-user MISO BC where the BS is equipped with *two* antennas. A fast fading channel without any time-correlation is considered, hence, no current CSIT can be predicted using the delayed CSIT. Receivers are assumed to have perfect instantaneous CSI and convey it back to the BS with a *one* time slot delay-link. Therefore, at time slot i, the BS has $\{\mathbf{h}_j(t)\}_{t=1}^{t=i-1}$ and Rx-j has $\{\mathbf{h}_j(t)\}_{t=1}^{t=i}$, where $\mathbf{h}_j(t) \in \mathbb{C}^{1\times 2}$ denotes the channel vector from the BS to Rx-j at time slot t. All the entries in the channel vectors are complex Gaussian variables. Moreover, as assumed in [54, 55], receivers have knowledge of all the beamforming vectors. Since this chapter is based on the MAT scheme, its brief introduction is given below.

3.2.1 The MAT Scheme for $\mathcal{S}_{MAT}^{(1)}$

Recall that in $\mathcal{S}_{MAT}^{(1)}$, there are *two* users and the BS is equipped with *two* antennas. The total scheme is made up of three time slots and two symbol vectors, $\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{C}^{2\times 1}$, are transmitted, where $\mathbf{s}_i = [s_i^{(1)}\ s_i^{(2)}]^T$ is for Rx-i. Briefly, the BS sends one of the symbol vectors to receivers in each of the first two time slots. Afterwards, the BS sends the sum of interference overheard by the two receivers in the first two time slots with the help of the delayed CSIT. In detail, at the first time slot, the BS sends the symbol vector $\mathbf{s}_1$. On one hand, Rx-1 regards it as desired signal but cannot decode it because only one observation of the two symbols is available. On the other hand, Rx-2 will treat the received signal as interference. Symmetrically, the BS sends $\mathbf{s}_2$ at the second time slot. At the beginning of the third time slot, the BS has acquired perfect knowledge of $\mathbf{h}_i(1)$ and $\mathbf{h}_i(2)$, where $i = 1, 2$, which enables it to reconstruct the sum of the overheard interference, i.e., $\mathbf{h}_2(1)\mathbf{s}_1 + \mathbf{h}_1(2)\mathbf{s}_2$. The symbols sent and received during the three time slots are shown in Figure 3.1.

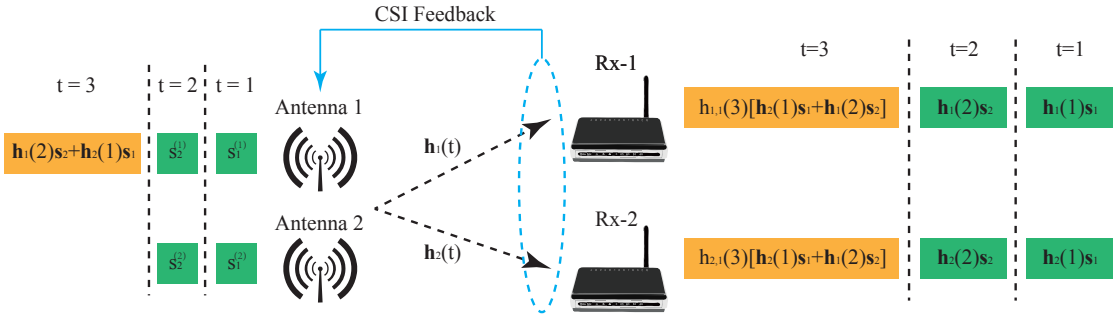


Figure 3.1: The MAT scheme for the 2-user MISO BC.

If each row represents one time slot, the signals received at the Rx-i are given by,

$$\mathbf{y}_i = \sqrt{\frac{P}{2}} \underbrace{\begin{bmatrix} \mathbf{h}_i(1) \\ \mathbf{0} \\ h_{i,1}(3)\mathbf{h}_2(1) \end{bmatrix}}_{\bar{\mathbf{H}}_{i1}} \mathbf{s}_1 + \sqrt{\frac{P}{2}} \underbrace{\begin{bmatrix} \mathbf{0} \\ \mathbf{h}_i(2) \\ h_{i,1}(3)\mathbf{h}_1(2) \end{bmatrix}}_{\bar{\mathbf{H}}_{i2}} \mathbf{s}_2 + \mathbf{n}_i,$$

where $h_{i,1}(3)$ is the channel entry from the first antenna at the BS to Rx-i at the third time slot and is only known by Rx-i. Symbol $\mathbf{n}_i = [n_i(1)\ n_i(2)\ n_i(3)]^T$ is the Gaussian noise vector at Rx-i with zero-mean and unit-variance. Symbol P denotes the total SNR at the BS. In this equation, the effective channel matrix for $\mathbf{s}_i$ at Rx-j is defined as $\bar{\mathbf{H}}_{ji}, j = 1, 2$. Clearly, $\bar{\mathbf{H}}_{11}$ and $\bar{\mathbf{H}}_{22}$ have rank *two* and $\bar{\mathbf{H}}_{12}$ and $\bar{\mathbf{H}}_{21}$ are rank *one* matrices. With this property, each receiver is able to obtain *two* observations of its *two* desired symbols and, meanwhile, eliminate the interference by making a subtraction of the two aligned interference signals. Specifically, the corresponding receiving matrices to suppress the interference at Rx-1 and Rx-2 are given by the following two equations, respectively:

$$\bar{\mathbf{U}}_1^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -h_{1,1}(3) & 1 \end{bmatrix} \quad \text{and} \quad \bar{\mathbf{U}}_2^T = \begin{bmatrix} 0 & 1 & 0 \\ -h_{2,1}(3) & 0 & 1 \end{bmatrix}.$$

For Rx-1, post-processing received signal is given by,

$$\hat{\mathbf{y}}_i = \bar{\mathbf{U}}_1^T \mathbf{y}_1 = \sqrt{\frac{P}{2}} \begin{bmatrix} \mathbf{h}_1(1) \\ \\ h_{1,1}(3)\mathbf{h}_2(1) \end{bmatrix} \mathbf{s}_1 + \begin{bmatrix} n_1(1) \\ \\ n_1(3) - h_{1,1}(3)n_1(2) \end{bmatrix},$$

where the two desired symbols contained in $\mathbf{s}_1$ can be decoded without interference. That is to say, *one* DoF can be obtained by decoding each desired symbol. Symmetrically, Rx-2 can decode its desired symbols by left multiplying its receiving matrix. In total, $\frac{4}{3}$ DoF are achieved per time slot.

3.2.2 The MAT Scheme for $\mathcal{S}_{MAT}^{(3)}$

The MAT scheme for $\mathcal{S}_{MAT}^{(3)}$ takes *three* phases consisting of *eight* time slots. There are *two* desired symbol vectors for each receiver, i.e., $\mathbf{s}_i \in \mathbb{C}^{2\times 1}$ and $\bar{\mathbf{s}}_i \in \mathbb{C}^{2\times 1}$ are for Rx-i, where $i = 1, 2, 3$. In the three time slots of the first phase, the BS sends $\mathbf{s}_1 + \mathbf{s}_2$, $\bar{\mathbf{s}}_1 + \mathbf{s}_3$ and $\bar{\mathbf{s}}_2 + \bar{\mathbf{s}}_3$, respectively. In each of the three time slots of the second phase, the BS sends the reconstructed interference overheard at the receivers in the first phase. For example, the overheard interference $\mathbf{h}_1(1)\mathbf{s}_2$ at Rx-1 and $\mathbf{h}_2(1)\mathbf{s}_1$ at Rx-2 will be sent at the first time slot of Phase 2. The third phase has two time slots in which two independent linear combinations of the interference overheard in the second phase are transmitted. A complete collection of the received symbols in the scheme are is given in Table 3.1. The common symbols sent in the third phase are given by,

$$c_i = u_i \underbrace{\mathbf{h}_3(4) \begin{bmatrix} \mathbf{h}_2(1)\mathbf{s}_1 \\ \\ \mathbf{h}_1(1)\mathbf{s}_2 \end{bmatrix}}_{y_3(4)} + v_i \underbrace{\mathbf{h}_2(5) \begin{bmatrix} \mathbf{h}_3(2)\bar{\mathbf{s}}_1 \\ \\ \mathbf{h}_1(2)\mathbf{s}_3 \end{bmatrix}}_{y_2(5)} + w_i \underbrace{\mathbf{h}_1(6) \begin{bmatrix} \mathbf{h}_3(3)\bar{\mathbf{s}}_2 \\ \\ \mathbf{h}_1(3)\bar{\mathbf{s}}_3 \end{bmatrix}}_{y_1(6)},$$

where $i = 1, 2$ and u_i, v_i, w_i are constants shared with the receivers. Note that $y_i(t)$ is the received signal at Rx-i at time slot t. When the receivers obtain symbols from all *eight* time slots, they will be able to decode the desired symbols using the backwards decoding approach as in [8]. Summing up, 12 symbols are sent within *eight* time slots, which means $\frac{3}{2}$ DoF are achieved in total per time slot.

		first time slot	second time slot	third time slot
Phase 1	Rx-1	$\mathbf{h}_1(1)(\mathbf{s}_1+\mathbf{s}_2)$	$\mathbf{h}_1(2)(\bar{\mathbf{s}}_1+\mathbf{s}_3)$	$\mathbf{h}_1(3)(\bar{\mathbf{s}}_2+\bar{\mathbf{s}}_3)$
	Rx-2	$\mathbf{h}_2(1)(\mathbf{s}_1+\mathbf{s}_2)$	$\mathbf{h}_2(2)(\bar{\mathbf{s}}_1+\mathbf{s}_3)$	$\mathbf{h}_2(3)(\bar{\mathbf{s}}_2+\bar{\mathbf{s}}_3)$
	Rx-3	$\mathbf{h}_3(1)(\mathbf{s}_1+\mathbf{s}_2)$	$\mathbf{h}_3(2)(\bar{\mathbf{s}}_1+\mathbf{s}_3)$	$\mathbf{h}_3(3)(\bar{\mathbf{s}}_2+\bar{\mathbf{s}}_3)$
Phase 2	Rx-1	$\mathbf{h}_1(4)\left[\mathbf{h}_2(1)\mathbf{s}_1\ \mathbf{h}_1(1)\mathbf{s}_2\right]^T$	$\mathbf{h}_1(5)\left[\mathbf{h}_3(2)\bar{\mathbf{s}}_1\ \mathbf{h}_1(2)\mathbf{s}_3\right]^T$	$\mathbf{h}_1(6)\left[\mathbf{h}_3(3)\bar{\mathbf{s}}_2\ \mathbf{h}_1(3)\bar{\mathbf{s}}_3\right]^T$
	Rx-2	$\mathbf{h}_2(4)\left[\mathbf{h}_2(1)\mathbf{s}_1\ \mathbf{h}_1(1)\mathbf{s}_2\right]^T$	$\mathbf{h}_2(5)\left[\mathbf{h}_3(2)\bar{\mathbf{s}}_1\ \mathbf{h}_1(2)\mathbf{s}_3\right]^T$	$\mathbf{h}_2(6)\left[\mathbf{h}_3(3)\bar{\mathbf{s}}_2\ \mathbf{h}_1(3)\bar{\mathbf{s}}_3\right]^T$
	Rx-3	$\mathbf{h}_3(4)\left[\mathbf{h}_2(1)\mathbf{s}_1\ \mathbf{h}_1(1)\mathbf{s}_2\right]^T$	$\mathbf{h}_3(5)\left[\mathbf{h}_3(2)\bar{\mathbf{s}}_1\ \mathbf{h}_1(2)\mathbf{s}_3\right]^T$	$\mathbf{h}_3(6)\left[\mathbf{h}_3(3)\bar{\mathbf{s}}_2\ \mathbf{h}_1(3)\bar{\mathbf{s}}_3\right]^T$
Phase 3	Rx-1	$h_{1,1}(7)c_1$	$h_{1,1}(8)c_2$	Empty
	Rx-2	$h_{2,1}(7)c_1$	$h_{2,1}(8)c_2$	
	Rx-3	$h_{3,1}(7)c_1$	$h_{3,1}(8)c_2$	

Table 3.1: Symbols received in all of the eight time slots of the 3-user MAT scheme with 2-antenna BS.

3.3 Probabilistic-Constrained Optimisation for $\mathcal{S}_{MAT}^{(1)}$

Although the MAT scheme is optimal from the DoF point of view, i.e., in the infinite SNR case, its performance can be improved in the finite SNR case. This is because, in the finite SNR case, the optimal beamforming vectors will not use all the DoF to eliminate the interference but will also consider the desired signal detectability. Therefore, after obtaining the delayed CSIT at the third time slot, instead of reconstructing the interference exactly the way they are, i.e., $\mathbf{h}_2(1)\mathbf{s}_1+\mathbf{h}_1(2)\mathbf{s}_2$, the BS can construct it as $\mathbf{w}_1\mathbf{s}_1+\mathbf{w}_2\mathbf{s}_2$, where $\mathbf{w}_i, i=1,2$ is the beamforming vector to be designed by the BS. In this way, the signals received during three time slots at Rx-i can be written as,

$$\mathbf{y}_i = \begin{bmatrix} y_i(1) \\ y_i(2) \\ y_i(3) \end{bmatrix} = \sqrt{\frac{P}{2}} \begin{bmatrix} \mathbf{h}_i(1) \\ \mathbf{0} \\ h_{i,1}(3)\mathbf{w}_1 \end{bmatrix} \mathbf{s}_1 + \sqrt{\frac{P}{2}} \begin{bmatrix} \mathbf{0} \\ \mathbf{h}_i(2) \\ h_{i,1}(3)\mathbf{w}_2 \end{bmatrix} \mathbf{s}_2 + \mathbf{n}_i, i, j = 1, 2, i \neq j,$$

where $y_i(t)$ denotes the signal received at time slot t.

After generalisation of the beamforming vectors in the MAT scheme, interference will not be perfectly aligned, i.e., interference leakage occurs. Therefore, the receivers can design their minimum interference leakage receiving matrices by generalising the receiving matrices in the MAT scheme into the following form,

$$\mathbf{U}_1^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -g_1 & 1 \end{bmatrix}, \quad \mathbf{U}_2^T = \begin{bmatrix} 0 & 1 & 0 \\ -g_2 & 0 & 1 \end{bmatrix},$$

where $g_i, i = 1, 2$ are complex scalars to be designed by the receivers. After multiplying the receiving matrices, the post-processing received signals at both receivers are given by (3.1) and (3.2), where the effective MIMO channels for Rx-1 have *zero* interference in the first data stream and $h_{1,1}(3)\mathbf{w}_2 - g_1\mathbf{h}_1(2)$ in the second data stream. Note that $\mathbf{G}_{ij}$ denotes the effective MIMO channel for the desired symbols or interference signal. To help better clarify our idea in the subsequent sections, the effective MIMO channel in the form of (3.1) and (3.2) is defined as *e-channel*. If $g_1 = h_{1,1}(3), g_2 = h_{2,1}(3)$ and $\mathbf{w}_1 = \mathbf{h}_2(1), \mathbf{w}_2 = \mathbf{h}_1(2)$, the scheme reduces to the original MAT scheme.

$$\begin{aligned}\hat{\mathbf{y}}_1 = \mathbf{U}_1^T\mathbf{y}_1 =& \sqrt{\frac{P}{2}}\underbrace{\begin{bmatrix}\mathbf{h}_1(1)\\ h_{1,1}(3)\mathbf{w}_1\end{bmatrix}}_{\mathbf{G}_{11}}\mathbf{s}_1 + \sqrt{\frac{P}{2}}\underbrace{\begin{bmatrix}\mathbf{0}\\ h_{1,1}(3)\mathbf{w}_2 - g_1\mathbf{h}_1(2)\end{bmatrix}}_{\mathbf{G}_{12}}\mathbf{s}_2 \\ &+ \begin{bmatrix}n_1(1)\\ n_1(3) - g_1 n_1(2)\end{bmatrix},\end{aligned} \tag{3.1}$$

$$\begin{aligned}\hat{\mathbf{y}}_2 = \mathbf{U}_2^T\mathbf{y}_2 =& \sqrt{\frac{P}{2}}\underbrace{\begin{bmatrix}\mathbf{h}_2(2)\\ h_{2,1}(3)\mathbf{w}_2\end{bmatrix}}_{\mathbf{G}_{22}}\mathbf{s}_2 + \sqrt{\frac{P}{2}}\underbrace{\begin{bmatrix}\mathbf{0}\\ h_{2,1}(3)\mathbf{w}_1 - g_2\mathbf{h}_2(1)\end{bmatrix}}_{\mathbf{G}_{21}}\mathbf{s}_1 \\ &+ \begin{bmatrix}n_2(2)\\ n_2(3) - g_2 n_2(1)\end{bmatrix},\end{aligned} \tag{3.2}$$

Theorem 1 *In the high-enough SNR regime (i.e., the SNR is high enough to generate tractable expression but small enough so that the effectiveness of desired signal detectability enhancement still exists) the achievable sum rate at receivers with the generalised MAT scheme can be approximated by the following equation:*

$$\begin{aligned}I(\mathbf{s}_1;\mathbf{y}_1) + I(\mathbf{s}_2;\mathbf{y}_2) \approx& \log\left(1 + \frac{P}{2}\|\mathbf{h}_1(1)\|^2\right) + \log\left(1 + \frac{P}{2}\|\mathbf{h}_2(2)\|^2\right) \\ &+ \log\left(\frac{\frac{P}{2}|h_{1,1}(3)|^2 + \frac{P^2}{4}|h_{1,1}(3)|^2\left(\|\mathbf{h}_1(1)\|^2 - \mathbf{w}_1\mathbf{h}_1^H(1)\mathbf{h}_1(1)\mathbf{w}_1^H\right)}{\left(1 + \frac{P}{2}\|\mathbf{h}_1(1)\|^2\right)\mathcal{Q}_2\hat{g}_2^2}\right) \\ &+ \log\left(\frac{\frac{P}{2}|h_{2,1}(3)|^2 + \frac{P^2}{4}|h_{2,1}(3)|^2\left(\|\mathbf{h}_2(2)\|^2 - \mathbf{w}_2\mathbf{h}_2^H(2)\mathbf{h}_2(2)\mathbf{w}_2^H\right)}{\left(1 + \frac{P}{2}\|\mathbf{h}_2(2)\|^2\right)\mathcal{Q}_1\hat{g}_1^2}\right).\end{aligned}$$

where $\|\mathbf{w}_1\| = \|\mathbf{w}_2\| = 1$ *is assumed and,*

$$\begin{aligned}\mathcal{Q}_1 &= 1 + \frac{P}{2}\|h_{1,1}(3)\mathbf{w}_2 - g_1\mathbf{h}_1(2)\|^2, \\ \mathcal{Q}_2 &= 1 + \frac{P}{2}\|h_{2,1}(3)\mathbf{w}_1 - g_2\mathbf{h}_2(1)\|^2, \\ \hat{g}_i^2 &= \max\{1, |g_i|^2\}, i = 1, 2.\end{aligned}$$

Proof: See Appendix A.

The closed-form expression in Theorem 1 provides a clear insight into the role of the beamforming vectors $\mathbf{w}_i$, where $i = 1, 2$, to the achievable sum rate. Specifically, by designing $\mathbf{w}_i$, the balance between the IA and the desired signals detectability enhancement can be struck. In order to

simplify the remaining description, the numerators containing only $\mathbf{w}_i$ are defined as $\Upsilon_i(\mathbf{w}_i)$ and the denominators containing only $\mathbf{w}_i$ are defined as $\Psi_i(\mathbf{w}_i)$, where $i = 1, 2$. Therefore,

$$\Upsilon_1(\mathbf{w}_1) = \frac{P}{2}|h_{1,1}(3)|^2 + \frac{P^2}{4}|h_{1,1}(3)|^2 \left(\|\mathbf{h}_1(1)\|^2 - \mathbf{w}_1\mathbf{h}_1^H(1)\mathbf{h}_1(1)\mathbf{w}_1^H\right), \tag{3.4a}$$

$$\Psi_1(\mathbf{w}_1) = \left(1 + \frac{P}{2}\|\mathbf{h}_1(1)\|^2\right)\mathcal{Q}_2\hat{g}_2^2,$$

$$\Upsilon_2(\mathbf{w}_2) = \frac{P}{2}|h_{2,1}(3)|^2 + \frac{P^2}{4}|h_{2,1}(3)|^2 \left(\|\mathbf{h}_2(2)\|^2 - \mathbf{w}_2\mathbf{h}_2^H(2)\mathbf{h}_2(2)\mathbf{w}_2^H\right),$$

$$\Psi_2(\mathbf{w}_2) = \left(1 + \frac{P}{2}\|\mathbf{h}_2(2)\|^2\right)\mathcal{Q}_1\hat{g}_1^2, \tag{3.4b}$$

where an important observation is that, in each logarithmic function in the expression of the achievable sum rate, the numerator and the denominator contain the same beamforming vector. Hence, in order to maximise the achievable sum rate, the logarithmic functions can be maximised independently. In other words, each of the beamforming vectors can be designed independently.

For the rest of the derivations, only the first logarithmic function is discussed in detail because the second one has a similar form. This logarithmic function cannot be optimised directly for two reasons: 1) It is not convex in terms of $\mathbf{w}_i$; 2) It contains unknown channel entries. Fortunately, both issues can be tackled at the same time using the probabilistic-constrained optimisation approach. The key idea is to keep a high probability that the interference leakage power is below a certain threshold and to maximise the signal detectability at the same time. In this way, $\|\mathbf{h}_1(1)\|^2 - \mathbf{w}_1\mathbf{h}_1^H(1)\mathbf{h}_1(1)\mathbf{w}_1^H$ is selected to be maximised and $\mathcal{Q}_2$ is kept a high probability that it is under a threshold. The former expression can be transformed with the following equation:

$$\begin{aligned}\|\mathbf{h}_1(1)\|^2 - \mathbf{w}_1\mathbf{h}_1^H(1)\mathbf{h}_1(1)\mathbf{w}_1^H &= \|\mathbf{h}_1(1)\|^2\mathbf{w}_1\mathbf{w}_1^H - \mathbf{w}_1\mathbf{h}_1^H(1)\mathbf{h}_1(1)\mathbf{w}_1^H \\ &= \mathbf{w}_1\left(\|\mathbf{h}_1(1)\|^2\mathbf{I}^{(2)} - \mathbf{h}_1^H(1)\mathbf{h}_1(1)\right)\mathbf{w}_1^H \\ &= \mathbf{w}_1\left(\mathbf{h}_1^H(1)\right)^{\perp}\mathbf{h}_1^{\perp}(1)\mathbf{w}_1^H,\end{aligned}$$

which is based on the fact that,

$$\left(\mathbf{h}_1^H(1)\right)^{\perp}\mathbf{h}_1^{\perp}(1) + \mathbf{h}_1^H(1)\mathbf{h}_1(1) = \|\mathbf{h}_1(1)\|^2\mathbf{I}^{(2)},$$

where $\mathbf{I}^{(j)}$ is the j-dimensional identity matrix. Therefore, the objective function of the optimisation problem can be written as,

$$\max_{\mathbf{w}_1} \quad \text{trace}\left(\mathbf{w}_1\left(\mathbf{h}_1^H(1)\right)^{\perp}\mathbf{h}_1^{\perp}(1)\mathbf{w}_1^H\right).$$

To guarantee a high probability that $\mathcal{Q}_2$ is below a certain threshold can be equivalently written as the following constraint:

$$\text{s.t.} \quad \Pr\left(\frac{P}{2}\|h_{2,1}(3)\mathbf{w}_1 - g_2\mathbf{h}_2(1)\|^2 \leq \bar{\alpha}\right) \geq \bar{\beta}, \tag{3.5}$$

where $0 \leq \bar{\beta} \leq 1$ and $\bar{\alpha}$ are parameters to be chosen according to the requirement of interference leakage power. Because of the difficulties in handling (3.5) directly, it is suitable to make the following transformation. By regarding g_2 as any complex number that is to be chosen, the norm $\|h_{2,1}(3)\mathbf{w}_1 - g_2\mathbf{h}_2(1)\|$ can be interpreted as the distance between any point in the line at the direction of $\mathbf{h}_2(1)$ and to the point $h_{2,1}(3)\mathbf{w}_1$. According to the property of complex space, the cosine of the angle ϑ between $\mathbf{w}_1$ and $\mathbf{h}_2(1)$ is $\cos(\vartheta) = \frac{\text{Re}(\mathbf{w}_1\mathbf{h}_2^H(1))}{\|\mathbf{w}_1\|\cdot\|\mathbf{h}_2(1)\|}$. Intuitively, the nearest distance from the direction $\mathbf{h}_2(1)$ to the point $h_{2,1}(3)\mathbf{w}_1$ is $\sin(\vartheta)|h_{2,1}(3)|\|\mathbf{w}_1\|$. An important notation of the latter property is that, because $\mathbf{w}_1$ has unit norm and $\mathbf{h}_2(1)$ can not be chosen, the distance is proportional to the norm of $h_{2,1}(3)$, which explains how does the probabilistic constraint works in the proposed scheme. Meanwhile, when $\mathbf{w}_1$ approaches the direction of $\mathbf{h}_2(1)$ in the complex space, the norm reduces and is equal to *zero* when $\mathbf{w}_1$ and $\mathbf{h}_2(1)$ are in the same direction. When $\mathbf{w}_1$ approaches $\mathbf{h}_2(1)$ it can be equivalently expressed as $\mathbf{w}_1$ being orthogonal to $\left(\mathbf{h}_2^H(1)\right)^\perp$ in the complex space. Therefore, the probabilistic constraint can be rewritten as,

$$\text{s.t.} \quad \Pr\left(\frac{P}{2}|h_{2,1}(3)|^2 \,\text{trace}\left(\mathbf{w}_1\left(\mathbf{h}_2^H(1)\right)^\perp \mathbf{h}_2^\perp(1)\mathbf{w}_1^H\right) \leq \hat{\alpha}\right) \geq \hat{\beta}.$$

where $\hat{\alpha}$ and $\hat{\beta}$ are also parameters to be chosen by the BS. In this way, the overall optimisation problem is formalised as,

$$\begin{aligned} \max_{\mathbf{w}_1} \quad & \text{trace}\left(\mathbf{w}_1\left(\mathbf{h}_1^H(1)\right)^\perp \mathbf{h}_1^\perp(1)\mathbf{w}_1^H\right) \\ \text{s.t.} \quad & \Pr\left(\frac{P}{2}|h_{2,1}(3)|^2 \,\text{trace}\left(\mathbf{w}_1\left(\mathbf{h}_2^H(1)\right)^\perp \mathbf{h}_2^\perp(1)\mathbf{w}_1^H\right) \leq \hat{\alpha}\right) \geq \hat{\beta} \\ & \|\mathbf{w}_1\|^2 = 1 \end{aligned} \tag{3.6a}$$

Theorem 2 *The probabilistic constraint in the probabilistic-constrained optimisation problem can be translated into the following inequality:*

$$\text{trace}\left(\mathbf{w}_1\left(\mathbf{h}_2^H(1)\right)^\perp \mathbf{h}_2^\perp(1)\mathbf{w}_1^H\right) \leq \frac{\gamma}{\sigma_h^2 P},$$

where each channel entry is $\mathcal{CN}(0, 2\sigma_h^2)$, and,

$$\gamma = \frac{\hat{\alpha}}{2\left(\text{erf}^{-1}(\sqrt{\hat{\beta}})\right)^2},$$

is the only parameter that is to be chosen by the BS.

Proof: See Appendix B.

Remark: Intuitively, with smaller γ, more DoF will be exploited to align the interference and with larger γ, more DoF can be exploited to enhance the desired signal detectability. Unfortunately, as in other probabilistic-constrained optimisation schemes [51, 52], the closed-form expression of the optimal γ that maximises the achievable sum rate cannot be obtained. However, a numerical approach will be provided in the simulation section, which shows that γ can be chosen straightforwardly without significant performance degradation. From another perspective, with a fixed γ, more DoF will be allocated to reduce the interference with larger P. In other words, in the infinite SNR case, this constraint makes the proposed optimisation approach reduce to the MAT scheme.

Based on Theorem 2, the optimisation problem to find $\mathbf{w}_1$ can be equivalently written as,

$$\max_{\mathbf{w}_1} \quad \text{trace}\left(\mathbf{w}_1 \left(\mathbf{h}_1^H(1)\right)^{\perp} \mathbf{h}_1^{\perp}(1)\mathbf{w}_1^H\right) \tag{3.7a}$$

$$\text{s.t.} \quad \text{trace}\left(\mathbf{w}_1 \left(\mathbf{h}_2^H(1)\right)^{\perp} \mathbf{h}_2^{\perp}(1)\mathbf{w}_1^H\right) \leq \frac{\gamma}{\sigma_h^2 P} \tag{3.7b}$$

$$\|\mathbf{w}_1\|^2 = 1 \tag{3.7c}$$

Note that the optimisation problem (3.7a)-(3.7c) is not convex. According to [56], it should be translated into the following equivalent optimisation problem:

$$\max_{\mathbf{W}_1} \quad \text{trace}\left(\mathbf{h}_1^{\perp}(1)\mathbf{W}_1 \left(\mathbf{h}_1^H(1)\right)^{\perp}\right) \tag{3.8a}$$

$$\text{s.t.} \quad \text{trace}\left(\mathbf{h}_2^{\perp}(1)\mathbf{W}_1 \left(\mathbf{h}_2^H(1)\right)^{\perp}\right) \leq \frac{\gamma}{\sigma_h^2 P} \tag{3.8b}$$

$$\text{trace}(\mathbf{W}_1) = 1 \tag{3.8c}$$

$$\mathbf{W}_1 \succcurlyeq \mathbf{0} \tag{3.8d}$$

where $\mathbf{W}_i = \mathbf{w}_i^H \mathbf{w}_i, i = 1, 2$ and $\mathbf{W}_i \succcurlyeq \mathbf{0}$ means that $\mathbf{W}_i$ is positive semi-definite matrices. As proven in [57], $\mathbf{W}_1$ will surely have rank one. Therefore, the singular vector of the solution of

(3.8a)-(3.8d) is equivalent to the solution of (3.7a)-(3.7c). As $\mathbf{W}_1$ is positive semi-definite, the optimisation problem (3.8a)-(3.8d) forms a semi-definite programming (SDP) problem and can be efficiently solved by optimisation tools such as the CVX tool in Matlab.

The two receivers can use the following optimisation problems to design their receiving matrices $\mathbf{U}_1$ and $\mathbf{U}_2$ so that the interference leakage power are minimised:

$$\min_{g_1 \in \mathbb{C}} \|\mathbf{G}_{12}\|_F^2, \qquad \min_{g_2 \in \mathbb{C}} \|\mathbf{G}_{21}\|_F^2,$$

which have the closed-form solution as,

$$g_1 = h_{1,1}(3)\mathbf{w}_2\mathbf{h}_1^H(2)\left(\mathbf{h}_1(2)\mathbf{h}_1^H(2)\right)^{-1}$$
$$g_2 = h_{2,1}(3)\mathbf{w}_1\mathbf{h}_2^H(1)\left(\mathbf{h}_2(1)\mathbf{h}_2^H(1)\right)^{-1}$$

By this stage, the iterative-free algorithm is completed.

3.4 Probabilistic-Constrained Optimisation for $\mathcal{S}_{MAT}^{(3)}$

For the case of 3-user and 3-antenna BS, the MAT scheme provides an optimal scheme that achieves the outer bound proposed in [8]. In this section, it will be shown that the optimal scheme, which takes *three* phases and *eight* time slots, can be generalised and decomposed into independent e-channels. Therefore, the proposed probabilistic-constrained optimisation approach can be utilised to maximise the achievable sum rate.

It can be seen from Table 3.1 that no beamforming is required in the first phase or the third phase. Note that Phase 3 is not an independent phase but only serves to help decode the symbols sent in Phase 2. In the second phase, beamforming technique is exploited when the BS sends the reconstructed interference observed by receivers at the corresponding time slot in the first phase. In this MAT scheme, the interference terms reconstructed and sent by the BS in the second phase can be written as,

$$\mathbf{x}_{MAT}(4) = \begin{bmatrix} \mathbf{h}_2(1)\mathbf{s}_1 \\ \\ \mathbf{h}_1(1)\mathbf{s}_2 \end{bmatrix}, \quad \mathbf{x}_{MAT}(5) = \begin{bmatrix} \mathbf{h}_3(2)\bar{\mathbf{s}}_1 \\ \\ \mathbf{h}_1(2)\mathbf{s}_3 \end{bmatrix}, \quad \mathbf{x}_{MAT}(6) = \begin{bmatrix} \mathbf{h}_3(3)\bar{\mathbf{s}}_2 \\ \\ \mathbf{h}_1(3)\bar{\mathbf{s}}_3 \end{bmatrix}. \tag{3.10}$$

Similar to (3.10) in Section 3.3, instead of reconstructing the interference terms exactly the way

$$\begin{bmatrix} y_i(1) \\ y_i(2) \\ y_i(3) \\ y_i(4) \\ y_i(5) \\ y_i(6) \end{bmatrix} = \sqrt{\frac{P}{2}} \begin{bmatrix} \mathbf{h}_i(1) \\ \mathbf{0} \\ \mathbf{0} \\ h_{i,1}(4)\mathbf{w}_1 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \mathbf{s}_1 + \sqrt{\frac{P}{2}} \begin{bmatrix} \mathbf{0} \\ \mathbf{h}_i(2) \\ \mathbf{0} \\ \mathbf{0} \\ h_{i,1}(5)\bar{\mathbf{w}}_1 \\ \mathbf{0} \end{bmatrix} \bar{\mathbf{s}}_1 + \sqrt{\frac{P}{2}} \begin{bmatrix} \mathbf{h}_i(1) \\ \mathbf{0} \\ \mathbf{0} \\ h_{i,2}(4)\mathbf{w}_2 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \mathbf{s}_2$$
$$+ \sqrt{\frac{P}{2}} \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{h}_i(3) \\ \mathbf{0} \\ \mathbf{0} \\ h_{i,1}(6)\bar{\mathbf{w}}_2 \end{bmatrix} \bar{\mathbf{s}}_2 + \sqrt{\frac{P}{2}} \begin{bmatrix} \mathbf{0} \\ \mathbf{h}_i(2) \\ \mathbf{0} \\ \mathbf{0} \\ h_{i,2}(5)\mathbf{w}_3 \\ \mathbf{0} \end{bmatrix} \mathbf{s}_3 + \sqrt{\frac{P}{2}} \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{h}_i(3) \\ \mathbf{0} \\ \mathbf{0} \\ h_{i,2}(6)\bar{\mathbf{w}}_3 \end{bmatrix} \bar{\mathbf{s}}_3 + \mathbf{n}_i \tag{3.11}$$

they are, they can be reconstructed as flexible combinations of symbols. In other words, only part of the DoF are exploited to align the interference. Firstly, the beamforming vectors are generalised and the interference terms are reconstructed as follows:

$$\mathbf{x}(4) = \begin{bmatrix} \mathbf{w}_1\mathbf{s}_1 \\ \\ \mathbf{w}_2\mathbf{s}_2 \end{bmatrix}, \quad \mathbf{x}(5) = \begin{bmatrix} \bar{\mathbf{w}}_1\bar{\mathbf{s}}_1 \\ \\ \mathbf{w}_3\mathbf{s}_3 \end{bmatrix}, \quad \mathbf{x}(6) = \begin{bmatrix} \bar{\mathbf{w}}_2\bar{\mathbf{s}}_2 \\ \\ \bar{\mathbf{w}}_3\bar{\mathbf{s}}_3 \end{bmatrix}.$$

The signals received in the first and second phases at Rx-i can be rewritten in a much more straightforward manner as in (3.11), Only the symbols received in the first two phases are listed, but note that Phase 3 can help to decode the symbols received in Phase 2, e.g., $h_{i,1}(4)\mathbf{w}_1\mathbf{s}_1$ and $h_{i,1}(4)\bar{\mathbf{w}}_1\bar{\mathbf{s}}_1$. Specifically, in the third phase of our proposed scheme, the BS sends two independent linear combinations of $y_3(4), y_2(5)$ and $y_1(6)$. Hence, with the knowledge of $y_1(6)$, which is the received signal at the sixth time slot at Rx-1, $y_3(4)$ and $y_2(5)$ can be successfully obtained as $y_3(4) = \mathbf{h}_3(4)\begin{bmatrix}\mathbf{w}_1\mathbf{s}_1 & \mathbf{w}_2\mathbf{s}_2\end{bmatrix}^T$ and $y_2(5) = \mathbf{h}_2(5)\begin{bmatrix}\bar{\mathbf{w}}_1\bar{\mathbf{s}}_1 & \mathbf{w}_3\mathbf{s}_3\end{bmatrix}^T$. By making use of the observations $y_1(4)$ and $y_1(5)$, the following equations can be obtained:

$$\begin{bmatrix} y_1(4) \\ \\ y_3(4) \end{bmatrix} = \sqrt{\frac{P}{2}} \begin{bmatrix} \mathbf{h}_1(4) \\ \\ \mathbf{h}_3(4) \end{bmatrix} \begin{bmatrix} \mathbf{w}_1\mathbf{s}_1 \\ \\ \mathbf{w}_2\mathbf{s}_2 \end{bmatrix}, \quad \begin{bmatrix} y_1(5) \\ \\ y_2(5) \end{bmatrix} = \sqrt{\frac{P}{2}} \begin{bmatrix} \mathbf{h}_1(5) \\ \\ \mathbf{h}_2(5) \end{bmatrix} \begin{bmatrix} \bar{\mathbf{w}}_1\bar{\mathbf{s}}_1 \\ \\ \mathbf{w}_3\mathbf{s}_3 \end{bmatrix},$$

With these two equations, Rx-1 can successfully decode $\mathbf{w}_1\mathbf{s}_1$ and $\bar{\mathbf{w}}_1\bar{\mathbf{s}}_1$.

The next task is to construct two linear combinations of symbols in $\mathbf{s}_1$ and $\bar{\mathbf{s}}_1$ which enables Rx-1 to decode them. With the direct observations of $\mathbf{s}_1$ and $\bar{\mathbf{s}}_1$ as well as their overheard interference terms from other receivers in hand, Rx-1 can successfully construct the following equation:

$$\begin{aligned}\begin{bmatrix} g_1 y_1(1) - h_{1,2}(4)\mathbf{w}_2\mathbf{s}_2 \\ y_1(4) - h_{1,2}(4)\mathbf{w}_2\mathbf{s}_2 \\ \bar{g}_1 y_1(2) - h_{1,2}(5)\mathbf{w}_3\mathbf{s}_3 \\ y_1(5) - h_{1,2}(5)\mathbf{w}_3\mathbf{s}_3 \end{bmatrix} &= \sqrt{\frac{P}{2}} \begin{bmatrix} g_1\mathbf{h}_1(1) & 0 \\ h_{1,1}(4)\mathbf{w}_1 & 0 \\ 0 & \bar{g}_1\mathbf{h}_1(2) \\ 0 & h_{1,1}(5)\bar{\mathbf{w}}_1 \end{bmatrix} \begin{bmatrix} \mathbf{s}_1 \\ \bar{\mathbf{s}}_1 \end{bmatrix} \\ &+ \sqrt{\frac{P}{2}} \begin{bmatrix} g_1\mathbf{h}_1(1) - h_{1,2}(4)\mathbf{w}_2 \\ 0 \\ 0 \\ 0 \end{bmatrix} \mathbf{s}_2 + \sqrt{\frac{P}{2}} \begin{bmatrix} 0 \\ 0 \\ \bar{g}_1\mathbf{h}_1(2) - h_{1,2}(5)\mathbf{w}_3 \\ 0 \end{bmatrix} \mathbf{s}_3,\end{aligned} \tag{3.12}$$

where the white noise is omitted. Recall that $g_1, \bar{g}_1$ are the complex scalars in the receiving matrices to be designed by Rx-1. It can be seen from (3.12) that Rx-1 can construct two independent e-channels to decode $\mathbf{s}_1$ and $\bar{\mathbf{s}}_1$. This means that, similar to Theorem 1, the achievable sum rate of $\mathbf{s}_1$ can be written as,

$$I(\mathbf{s}_1; y_1(1), y_1(4), \mathbf{w}_2\mathbf{s}_2) = \mathcal{Q}_4 + \log\left(\frac{\Upsilon_1(\mathbf{w}_1)}{\Psi_2(\mathbf{w}_2)}\right),$$

where $\mathcal{Q}_4$ stands for all the terms independent of $\mathbf{w}_i$. Symbols $\Upsilon_i(\mathbf{w}_i)$ and $\Psi_i(\mathbf{w}_i)$ represent functions related to the beamforming vector $\mathbf{w}_i$ and are independent of any another $\mathbf{w}_j, j \neq i$. Because the effective MIMO channel used to decode the private symbols is the same as that in the 2-user case, the specific realisations of $\Upsilon_i(\cdot)$ and $\Psi_i(\cdot)$ can be found with similar methods as obtaining (3.4a)-(3.4b). The sum of the achievable sum rate to decode all of the desired symbols is given by,

$$\begin{aligned}I_{\sum} = \mathcal{Q}_4 &+ \log\left(\frac{\Upsilon_1(\mathbf{w}_1)}{\Psi_2(\mathbf{w}_2)}\right) + \log\left(\frac{\bar{\Upsilon}_1(\bar{\mathbf{w}}_1)}{\Psi_3(\mathbf{w}_3)}\right) + \log\left(\frac{\Upsilon_2(\mathbf{w}_2)}{\Psi_1(\mathbf{w}_1)}\right) \\ &+ \log\left(\frac{\bar{\Upsilon}_2(\bar{\mathbf{w}}_2)}{\bar{\Psi}_3(\bar{\mathbf{w}}_3)}\right) + \log\left(\frac{\Upsilon_3(\mathbf{w}_3)}{\bar{\Psi}_1(\bar{\mathbf{w}}_1)}\right) + \log\left(\frac{\bar{\Upsilon}_3(\bar{\mathbf{w}}_3)}{\bar{\Psi}_2(\bar{\mathbf{w}}_2)}\right)\end{aligned}$$

$$\begin{aligned}= \mathcal{Q}_4 &+ \log\left(\frac{\Upsilon_1(\mathbf{w}_1)}{\Psi_1(\mathbf{w}_1)}\right) + \log\left(\frac{\bar{\Upsilon}_1(\bar{\mathbf{w}}_1)}{\bar{\Psi}_1(\bar{\mathbf{w}}_1)}\right) + \log\left(\frac{\Upsilon_2(\mathbf{w}_2)}{\Psi_2(\mathbf{w}_2)}\right) \\ &+ \log\left(\frac{\bar{\Upsilon}_2(\bar{\mathbf{w}}_2)}{\bar{\Psi}_2(\bar{\mathbf{w}}_2)}\right) + \log\left(\frac{\Upsilon_3(\mathbf{w}_3)}{\Psi_3(\mathbf{w}_3)}\right) + \log\left(\frac{\bar{\Upsilon}_3(\bar{\mathbf{w}}_3)}{\bar{\Psi}_3(\bar{\mathbf{w}}_3)}\right),\end{aligned} \tag{3.13}$$

where $\bar{\Upsilon}_i(\bar{\mathbf{w}}_i)$, $\bar{\Psi}_i(\bar{\mathbf{w}}_i)$ have similar definition to $\Upsilon_i(\mathbf{w}_i)$, $\Psi_i(\mathbf{w}_i)$. Because each logarithmic function in (3.13) only contains one beamforming vector, maximising each logarithmic function independently can lead to the maximisation of the sum rate. For example, to maximise $\log\left(\frac{\Upsilon_1(\mathbf{w}_1)}{\Psi_1(\mathbf{w}_1)}\right)$, the probabilistic-constrained optimisation problem is formalised in a way similar to (3.8a)-(3.8d) as follows:

$$\begin{aligned}\max_{\mathbf{W}_1} \quad & \text{trace}\left(\mathbf{h}_1^{\perp}(1)\mathbf{W}_1\left(\mathbf{h}_1^H(1)\right)^{\perp}\right) \\ \text{s.t.} \quad & \text{trace}\left(\mathbf{h}_2^{\perp}(1)\mathbf{W}_1\left(\mathbf{h}_2^H(1)\right)^{\perp}\right) \leq \frac{\gamma}{\sigma_h^2 P} \\ & \text{trace}(\mathbf{W}_1) = 1 \\ & \mathbf{W}_1 \succcurlyeq \mathbf{0}\end{aligned}$$

where $\mathbf{w}_1$ is the eigenvector of $\mathbf{W}_1$. Rx-1 can minimise the interference leakage power by setting the variables in (3.12) using the following equation:

$$\begin{aligned}g_1 =& h_{1,2}(4)\mathbf{w}_2\mathbf{h}_1^H(1)\left(\mathbf{h}_1(1)\mathbf{h}_1^H(1)\right)^{-1}, \\ \bar{g}_1 =& h_{1,2}(5)\mathbf{w}_3\mathbf{h}_1^H(2)\left(\mathbf{h}_1(2)\mathbf{h}_1^H(2)\right)^{-1}.\end{aligned}$$

All of the other parameters g_i and $\bar{g}_i$ where $i = 2, 3$ can be found using a similar method.

3.5 Simulation Results

In order to realise the full potential of the proposed probabilistic-constrained optimisation method, a spatial correlation at the receiver side is considered. Note that a spatial correlation is also considered in the simulations of [45]. The reason behind this is that, in the presence of a spatial correlation at the receiver side, spending all of the DoF to align the interference tends to be more suboptimal because this will cause a correlation between the two observations of the desired symbols. Therefore, in the presence of a large spatial correlation at the receiver side, enhancing the desired signal detectability becomes more effective. The correlated channel can be modelled

as follows:

$$\mathbf{H} = \mathbf{R}_r^{1/2}\mathbf{H}_w,$$

where $\mathbf{R}_r$ is the receiver correlation matrices, with diagonal elements being *one* and others being r_r, where $r_r \in [0, 1)$. Matrix $\mathbf{H}_w$ denotes the channel with i.i.d. $\mathcal{CN}(0, 1)$ entries, i.e., $\sigma_h = 2^{-\frac{1}{2}}$. Note that each row of $\mathbf{H}$ is regarded as the channel vector from the BS to each of the receivers. To more clearly show the sum rate gain, a relatively large spatial correlation is assumed, i.e., $r_r = 0.8$, unless stated otherwise. Note that r_r can be calculated using the antenna spacing, the angle of arrival and the angular spread by the formulas derived in [58], but it is out of the scope of this thesis. The beamforming vectors of the dual SINR scheme are obtained as the closed-form solution in [45] (i.e., Eq.(70) and Eq.(71)). The beamforming vectors for the MAT scheme simply assume that $\mathbf{w}_1 = \mathbf{h}_2(1)$, $\mathbf{w}_1 = \mathbf{h}_1(2)$ with the power constraint as $\|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 = 2$. Each point in the simulation results is averaged over 500 channel realisations.

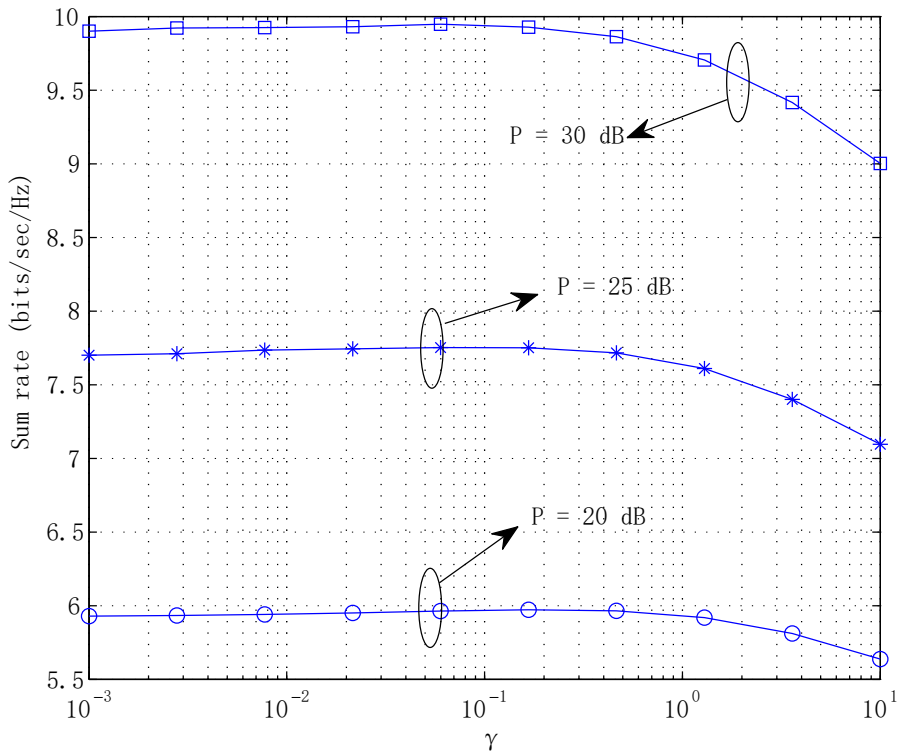


Figure 3.2: Sum rate for 2-user MISO BC versus γ with various SNRs.

Figure 3.2 demonstrates the achievable sum rate versus γ for various SNRs, i.e., $P = 20$ dB, $P = 25$ dB and $P = 30$ dB. The major motivation of this simulation is to provide a numerical basis from which to select γ. When $\gamma \to 0$, all of the DoF are exploited to align the interference. Thus, the proposed scheme reduces to the MAT scheme (according to the numerical simulation,

the MAT scheme witnesses an interference leakage power at the level of $\gamma = 10^{-32}$). When γ increases, more DoF are used to increase the desired signal orthogonality. On the one hand, it can be seen that the lines with different SNRs achieve the highest sum rates at approximately the same γ. On the other hand, below the optimal points, the sum rates increase quite slowly with γ. Therefore, a small-enough γ (i.e., smaller than the optimal value but large enough so that it still outperforms the MAT scheme considerably) can be chosen with insignificant sum rate degradation. For example, choosing $\gamma = 0.001$ will cause less than 0.1 bits/s/Hz of sum rate degradation compared to the highest value. Based on this observation, it can be concluded that striking the balancing between aligning interference and increasing the signal detectability as in [45] will not significantly outperform our proposed scheme with a small-enough γ value.

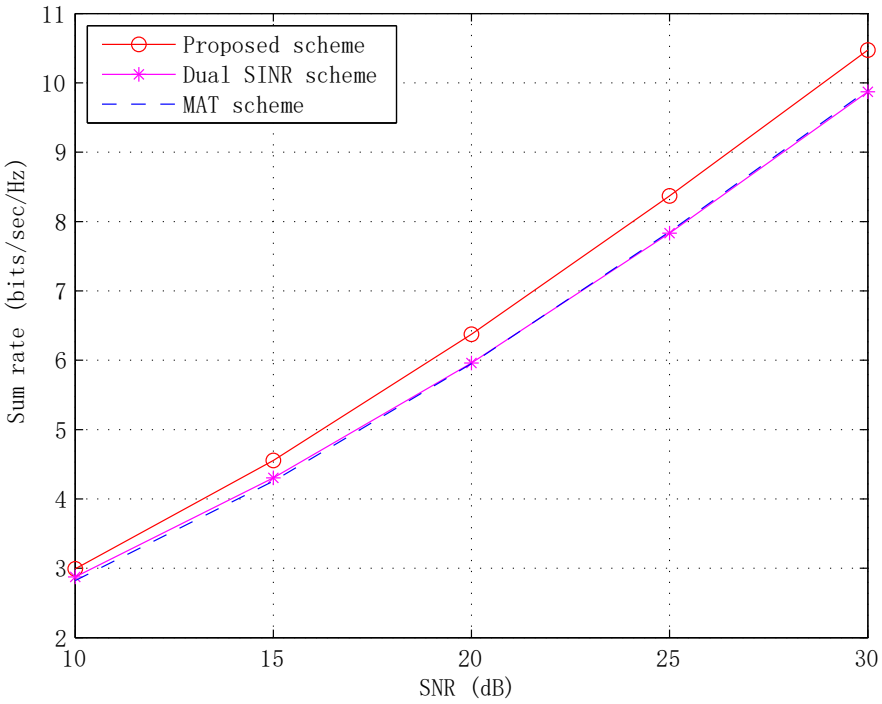


Figure 3.3: Sum rate of 2-user MISO BC with a large spatial correlation at the receiver side.

The effectiveness of the proposed scheme can be seen in Figure 3.3 where sum rates achieved by the MAT scheme, the dual SINR scheme [45] and the proposed probabilistic-constrained optimisation method are depicted for the 2-user case. Based on the first simulation, $\gamma = 0.1$ can be safely chosen with negligible sum rate degradation. One important observation is that all three lines tend to have the same slope in the high SNR region, which indicates that all of them achieve the same DoF, as expected. From another perspective, when the SNR is low, the constraint derived in Theorem 2 guides the proposed probabilistic-constrained optimisation

method to allocate more DoF to enhance the desired signal detectability. Thus, the performance gain increases. In the high SNR region, this constraint makes our proposed scheme reduce to the MAT scheme, which is optimal in this scenario. Thus, the performance gain does not continuously increase. Note that the achievable sum rate derived in Theorem 1 is only valid for a high-enough SNR scenario, which explains why in the low SNR region, e.g., $P = 10$ dB, the performance gain is not substantial. Another interesting observation is that the dual SINR line almost overlaps with the MAT scheme line, which is different from the simulation results in [45]. This is mainly because of the receiver structure mismatch. Specifically, the receiver structure is circumvented in [45], but minimum interference leakage receivers are considered in this chapter. This mismatch offsets the gain from effectively compromising between aligning the interference and enhancing the desired signal detectability.

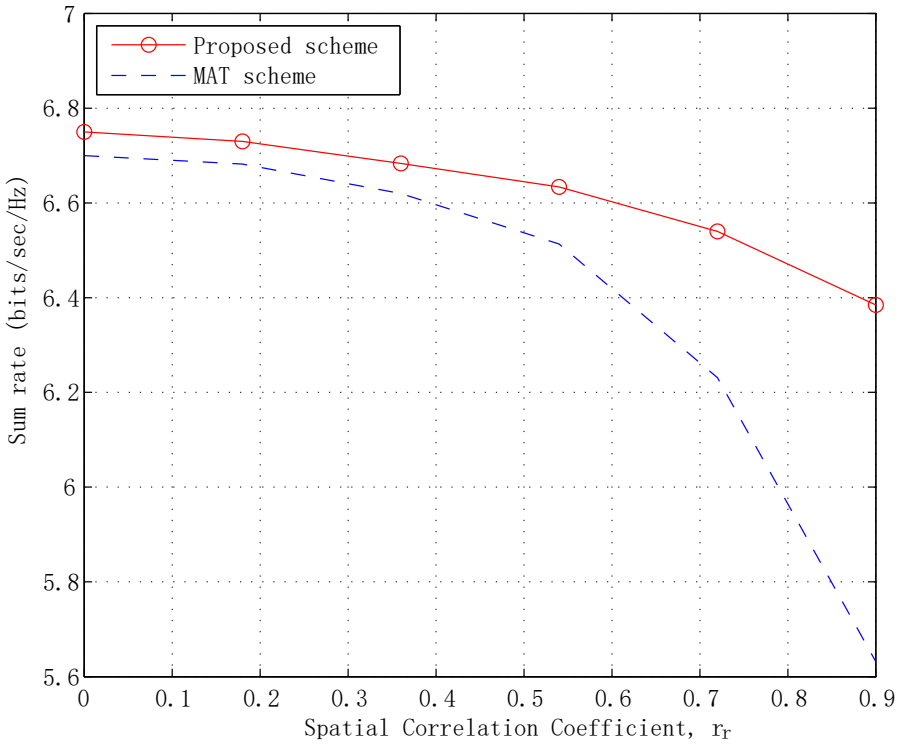


Figure 3.4: Sum rate of 2-user MISO BC versus the spatial correlation at the receiver side.

Figure 3.4 depicts the sum rate versus the spatial correlation at the receiver side, i.e., r_r, in order to better understand their relationship. The condition $P = 20$ dB and $\gamma = 0.001$ is considered in this simulation. Note that, although the accurate relationship between r_r and the optimal γ is out of the scope of this chapter, a near-optimal solution can be found using the numerical results observed from Figure 3.2, i.e., select a small-enough γ. In practice, proper γ values for various

r_r can be found through Monte Carlo simulations. However, choosing $\gamma = 0.001$ is enough to deliver the key point for this simulation. The most important observation from Figure 3.4 is that the performance gain increases with r_r. This observation validates our conjecture that a larger spatial correlation at the receiver side will cause a larger correlation between the two observations of the desired symbols and, therefore, enhancing the desired signal detectability becomes more effective.

3.6 Summary

In this work, an iterative-free probabilistic-constrained optimisation method was proposed to enhance the achievable sum rate of the MAT scheme in the finite SNR case. MISO BC with 2-antenna BS in the presence of perfect delayed CSIT was considered. At first, the beamforming vectors and the receiving matrices in the MAT scheme were generalised, which allowed them to allocate part of the DoF to enhance the desired signal detectability while cancelling the interference. Next, while enhancing the desired signal detectability, the generalised beamforming vectors were designed to guarantee a high probability that the interference leakage power at the minimum interference leakage receivers would be below a certain threshold. Notably, the probabilistic constraints were formalised by taking the distribution of unknown channel entries into consideration. At the receiver side, the closed-form solution for the receiving matrices was provided. This probabilistic-constrained optimisation method was first exploited to the MAT scheme in $\mathcal{S}_{MAT}^{(1)}$. Afterwards, it was illustrated that the effective MIMO channel of the MAT scheme in $\mathcal{S}_{MAT}^{(3)}$ can be decomposed into multiple effective MIMO channels of the MAT scheme in $\mathcal{S}_{MAT}^{(1)}$. Thus, the achievable sum rate by the MAT scheme for both settings could be enhanced by the proposed probabilistic-constrained optimisation approach. Simulation results showed that the sum rate was not sensitive to the allowed interference leakage power factor γ once it was kept under the optimal point. In other words, selecting a small-enough γ would cause negligible sum rate degradation. It could also be witnessed that the proposed probabilistic-constrained beamforming approach outperformed the MAT scheme and the dual SINR scheme in the presence of minimum interference leakage receivers especially with a large spatial correlation at the receiver side. Note that the performance gain tended to be insignificant with a small spatial correlation at the receiver side.

Chapter 4

Degrees of Freedom Regions Using Delayed CSIT in The Time-Correlated Channels

4.1 Introduction

The MAT scheme challenges the commonly accepted viewpoint that delayed CSIT will not bring any DoF gain and validates its usefulness. For example, in Section 3.2, how the MAT scheme achieves $\frac{4}{3}$ DoF in the 2-user MISO system using only delayed CSIT was introduced, and the common knowledge is that only *one* DoF can be achieved if no current CSIT is available. Although inspiring in nature, the MAT scheme relies somewhat on the overly pessimistic assumption that no current CSI is available at the BS. In the case when the channel coherence time is large compared to the CSI feedback delays, the obtained CSIT will be correlated to the current channel, i.e., the delayed CSIT can be exploited to predict the current CSIT. Now that the MAT has established the usefulness of the delayed CSIT, intuitively, the combination of the delayed CSIT and the current CSIT can be exploited to further enhance the achievable DoF. Therefore, this chapter is devoted to study the DoF regions and the corresponding achievability schemes using delayed CSIT in the presence of a time-correlation in the channel.

Recall that in Section 2.2.2, the differences between BC and IC was introduced, where the BS in BC always knows all the private symbols but the transmitters in IC only know the private symbols for their intended receiver. Having this in mind, it is interesting to investigate whether this difference influences the DoF region achieved by IC and BC with the same number of antennas and CSIT availability. Hence, in this chapter, it is shown that the schemes based on the MAT schemes in $\mathcal{S}_{MAT}^{(1)}$ can be extended to IC and exactly the same DoF can be achieved. Interestingly, extending the schemes based on the MAT scheme in $\mathcal{S}_{MAT}^{(2)}$ and $\mathcal{S}_{MAT}^{(3)}$ to IC is a challenging task because special combinations of all the private symbols are required. More specific literature reviews and dedicated objectives are provided in each section.

The remainder of this chapter is organised as follows. Section 4.2 provides the common system model for the subsequent sections. Section 4.3 considers $\mathcal{S}_{MAT}^{(1)}$ with imperfect current and delayed CSIT. Afterwards, an improved scheme for the same CSIT availability and system setting is provided in Section 4.4. Section 4.5 considers $\mathcal{S}_{MAT}^{(2)}$ with perfect delayed CSIT as well as imperfect current CSIT. The same CSIT availability in $\mathcal{S}_{MAT}^{(3)}$ is considered in Section 4.6. Simulation results are given in Section 4.7 and conclusive remarks can be found in Section 4.8.

4.2 System Model

In this chapter, both BC and IC are considered wherein the transmitters are equipped with M antennas and receivers are equipped with N antennas. The CSIT under consideration is exploited in terms of its current and delayed knowledge subject to realistic conditions, assuming that it is imperfect with different degrees of imperfection between current and delayed CSIT. Actually, current CSIT is received during the coherence period of the channel, while delayed CSIT is obtained after the end of the coherence period. Moreover, current CSIT is estimated by its correlation with delayed CSIT; therefore, current CSIT is subject to more channel estimation error compared to delayed CSIT. For instance, in the MISO IC, the current channel from Tx-j to Rx-i at time slot t, denoted as $\mathbf{h}_{ij}(t)$, is estimated by Tx-i as $\hat{\mathbf{h}}_{ij}(t)$ using the following equation:

$$\mathbf{h}_{ij}(t) = \hat{\mathbf{h}}_{ij}(t) + \tilde{\mathbf{h}}_{ij}(t), \tag{4.1}$$

where the estimation error matrix $\tilde{\mathbf{h}}_{ij}(t)$ is composed by Gaussian entries with *zero* mean and variance $P^{-\alpha}$. Note that $\tilde{\mathbf{h}}_{ij}(t)$ is independent from $\mathbf{h}_{ij}(t)$ and $\hat{\mathbf{h}}_{ij}(t)$. The parameter α describes the quality of the estimated channel at high SNR. Thus, $\alpha = 0$ indicates no current CSIT, while $\alpha \to \infty$ denotes perfect instantaneous CSIT. However, when $\alpha \geq 1$, the ZF beamforming approach may itself achieve negligible DoF loss due to sufficient knowledge of current CSIT [26]. On that account, only $\alpha \in [0, 1]$ will be considered in this chapter.

As far as the imperfect delayed CSIT is concerned, transmitters can obtain the delayed estimates of the real channel matrices at the end of the time slot with the feedbacks from receivers. The delayed CSIT can be modelled as,

$$\mathbf{h}_{ij}(t) = \check{\mathbf{h}}_{ij}(t) + \ddot{\mathbf{h}}_{ij}(t), \tag{4.2}$$

where $\check{\mathbf{h}}_{ij}(t)$ and $\ddot{\mathbf{h}}_{ij}(t)$ are the estimated delayed CSIT and the accompanied estimation error.

The error vector is independent from $\mathbf{h}_{ij}(t)$ and $\check{\mathbf{h}}_{ij}(t)$. In fact, the estimation error is modelled similarly to the current estimation error given above, i.e., it is described by means of an exponential function of the SNR [59–61]. Hence, if the accuracy of the delayed channel is denoted as β, the variance of the delayed estimation error of each element is $P^{-\beta}$. No CSIT available can be expressed as $\beta = 0$ while perfect delayed CSIT available can be expressed $\beta = 1$ by following a rationale similar to the case of the current estimation error. Because current CSIT is subject to more estimation error compared to delayed CSIT, $0 \leq \alpha \leq \beta \leq 1$ is considered in this chapter.

In this chapter, $\mathbf{h}_{ij}$, $\mathbf{H}_{ij}$, $\mathbf{H}_{ij}(s,t)$, $\mathbf{H}_i$ and $\mathbf{H}_i(t)$ also denote the channel vectors or matrices just like $\mathbf{h}_{ij}(t)$. Specifically,

- $\mathbf{h}_{ij}$ denotes the channel vector from Tx-j to Rx-i;
- $\mathbf{H}_{ij}$ denotes the channel matrix from Tx-j to Rx-i;
- $\mathbf{H}_{ij}(s,t)$ denotes the channel matrix from Tx-j to Rx-i at time slot t of Phase s;
- $\mathbf{H}_i$ denotes the channel matrix from the BS to Rx-i;
- $\mathbf{H}_i(t)$ denotes the channel matrix from the BS to Rx-i at time slot t.

The corresponding current and delayed estimates of the above five symbols have similar definitions as in (4.1) and (4.2). Note that all of the private symbols are mutually independent across schemes. Finally, this chapter follows two usual conventions [54,59,60,62,63]: 1) Receivers are assumed to have perfect and global CSI as well as all the corresponding estimates at the transmitters. 2) The transmitters in the IC have current and delayed estimates of the links to which they are connected.

4.3 Scheme for $\mathcal{S}^{(1)}_{MAT}$ with Imperfect Current and Delayed CSIT

The authors in [62] studied the DoF region in $\mathcal{S}^{(1)}_{MAT}$, i.e., $K = 2$, $M = 2$ and $N = 1$, with imperfect current CSIT as well as perfect delayed CSIT, where each receiver is able to achieve $\frac{2+\alpha}{3}$ DoF. The authors also predicted that, if the delayed CSIT was also imperfect, $\frac{1+\alpha+\beta}{3}$ DoF would be achieved by each receiver. However, no specific achievability scheme was provided to achieve such DoF. Meanwhile, the authors in [64] showed that the schemes in [62] could be extended to IC and exactly the same DoF could be achieved. Based on the extension in [64] and the prediction from [62], theoretically, the scheme for $\mathcal{S}^{(1)}_{MAT}$ IC with imperfect current and delayed CSIT should also achieve $\frac{1+\alpha+\beta}{3}$ DoF per receiver. Thus, the objective of this section is

to achieve such DoF and show that the scheme can be extended to BC.

4.3.1 Degrees of Freedom Region

Theorem 3 *The DoF region achieved by the MISO IC with both imperfect delayed and current CSIT is given by,*

$$\begin{aligned} d_1 \leq 1, \quad d_2 \leq 1, \\ (1+\beta-2\alpha)d_1 + (2-\alpha-\beta)d_2 \leq (1-\alpha)(1+\beta+\alpha), \\ (2-\alpha-\beta)d_1 + (1+\beta-2\alpha)d_2 \leq (1-\alpha)(1+\beta+\alpha), \end{aligned}$$

where d_i is the DoF achieved by Rx-i.

The DoF region is shown in Figure 4.1 and forms a polygon with the following vertices:

$$\left\{(0,0),(1,0),(1,\alpha),\left(\frac{1+\alpha+\beta}{3},\frac{1+\alpha+\beta}{3}\right),(\alpha,1),(0,1)\right\}.$$

Remark: When perfect delayed CSI is available, i.e. $\beta = 1$, the symmetric point becomes $(\frac{2+\alpha}{3}, \frac{2+\alpha}{3})$, which reduces to [64]. The point $(1,1)$ is associated with both perfect current and delayed CSIT, i.e. $\alpha = \beta = 1$. The line linking $(1,0)$ and $(0,1)$ is by TDMA. When α and β both tend to be *zero*, DoF achieved by the proposed scheme are less than TDMA. This is due to the proposed scheme have *three* time slots while the TDMA is considered in *two* time slots. With small α and β, the extra time slot can not contribute to the DoF gain.

4.3.2 Scheme $\mathcal{X}_1$ for Vertex $M^{(1)}$

The corresponding scheme is divided into three time slots.

Time Slot 1

In order to exploit the spatial modulation, the symbols are beamformed as the following equations:

$$\begin{aligned} \mathbf{x}_1(1) &= \left[\mathbf{q}_{21}(1) \quad \mathbf{q}_{21}^{\perp}(1)\right]\mathbf{s}_1, \\ \mathbf{x}_2(1) &= \left[\mathbf{q}_{12}(1) \quad \mathbf{q}_{12}^{\perp}(1)\right]\mathbf{s}_2, \end{aligned}$$

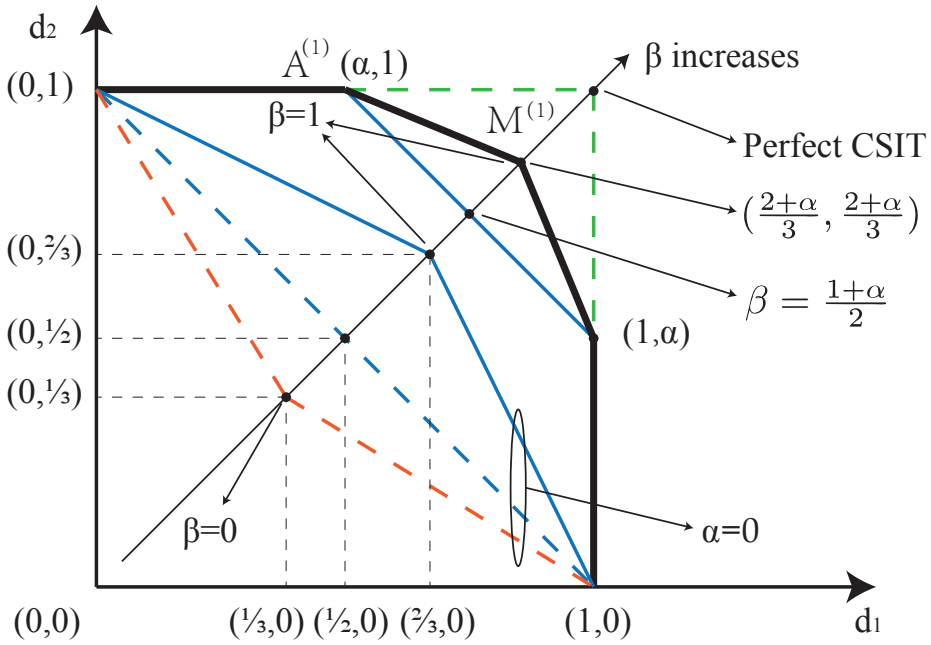


Figure 4.1: The DoF region of 2-user MISO IC using the basic scheme.

where the private symbol vector from Tx-i is $\mathbf{s}_i = [s_i^{(1)}\ s_i^{(2)}]^T$ where $\mathbb{E}[|s_1^{(1)}|^2] \doteq \mathbb{E}[|s_2^{(1)}|^2] \doteq P^{1-\alpha}$ and $\mathbb{E}[|s_1^{(2)}|^2] \doteq \mathbb{E}[|s_2^{(2)}|^2] \doteq P$. Symbol $\mathbf{q}_{ij}(t)$ is the normalised beamforming vector which lies in the range of $\hat{\mathbf{h}}_{ij}(t)$ and $\mathbf{q}_{ij}^{\perp}(t)$ is the normalised beamforming vector lying in the orthogonal space of $\hat{\mathbf{h}}_{ij}(t)$. The symbol $\doteq$ means the two terms increase at the same rate when $P \to \infty$. The corresponding received signals at Rx-1 and Rx-2 are given by,

$$y_1(1) = \mathbf{h}_{11}(1)\mathbf{x}_1(1) + \overbrace{\mathbf{h}_{12}(1)\mathbf{x}_2(1)}^{\eta_1} + n_1(1),$$
$$y_2(1) = \mathbf{h}_{22}(1)\mathbf{x}_2(1) + \overbrace{\mathbf{h}_{21}(1)\mathbf{x}_1(1)}^{\eta_2} + n_2(1),$$

where η_1 and η_2 are the overheard interference terms at Rx-1 and Rx-2, respectively, and $n_i(t)$ is the zero-mean unit-variance additive white noise. Note that the over-brace in this chapter means redefine the included term(s) with a new symbol. The interference overheard at Rx-1, i.e., η_1, can be decomposed into the following terms:

$$\eta_1 = \mathbf{h}_{12}(1)\mathbf{x}_2(1) = \mathbf{h}_{12}(1)\mathbf{q}_{12}(1)\mathbf{s}_2^{(1)} + \mathbf{h}_{12}(1)\mathbf{q}_{12}^{\perp}(1)\mathbf{s}_2^{(2)}$$
$$= \overbrace{\underbrace{\check{\mathbf{h}}_{12}(1)\mathbf{q}_{12}(1)s_2^{(1)}}_{P^{1-\alpha}} + \underbrace{\check{\mathbf{h}}_{12}(1)\mathbf{q}_{12}^{\perp}(1)s_2^{(2)}}_{P^{1-\alpha}}}^{\check{\eta}_1} + \overbrace{\underbrace{\ddot{\mathbf{h}}_{12}(1)\mathbf{q}_{12}(1)s_2^{(1)}}_{P^{1-\alpha-\beta}} + \underbrace{\ddot{\mathbf{h}}_{12}(1)\mathbf{q}_{12}^{\perp}(1)s_2^{(2)}}_{P^{1-\beta}}}^{\ddot{\eta}_1},$$

where symbol in the under-brace represents the power of this term in the infinite SNR case. The first term $\check{\eta}_1$ is the part of interference overheard by Rx-1 that can be reconstructed based on the delayed CSIT at Tx-2. Accordingly, $\ddot{\eta}_1$ is the part of the interference that can not be reconstructed because of the imperfection of the delayed CSIT, i.e., the estimation error of the interference. Because of the symmetry of the scheme, the power of the reconstructed interference terms at Rx-1 and Rx-2 is $\mathbb{E}[|\check{\eta}_1|^2] \doteq \mathbb{E}[|\check{\eta}_2|^2] \doteq P^{1-\alpha}$ and $\mathbb{E}[|\ddot{\eta}_1|^2] \doteq \mathbb{E}[|\ddot{\eta}_1|^2] \doteq P^{1-\beta}$.

Instead of transmitting the overheard interference terms in an analog fashion as usual, interference terms is digitised using the following equation:

$$\check{\eta}_i = \bar{\check{\eta}}_i + \tilde{\eta}_i,$$

where $\bar{\check{\eta}}_i$ and $\tilde{\eta}_i$ denote the digitised value and the digitisation error, respectively, of the interference η_i. The reason behind this is to suppress the interference power and make room for new private symbols. Especially when current CSIT quality is high, which leads to low interference power, there will be mismatch between reconstructed interference power and allowable transmit power.

The digitisation rate is chosen in order to drown the digitisation error in the white noise so that the interference can be decoded successfully. The rate distortion theory states that in order to emerge the digitisation error $\tilde{\eta}_i$ in the white noise, the number of bits required in $\bar{\check{\eta}}_i$ is given by,

$$\mathbb{E}\left[\log\left(\frac{\sigma^2_{\check{\eta}_i}}{\sigma^2_D}\right)\right] \text{ bits.}$$

where $\sigma^2_{\check{\eta}_i}$ is the variance of $\check{\eta}_i$ and σ^2_D is the required variance of digitisation error [65]. Since the power of the reconstructed interference is $\mathbb{E}[|\check{\eta}_i|^2] \doteq P^{1-\alpha}$, to get $\sigma^2_D = 1$, $\mathbb{E}\left[\log \sigma^2_{\check{\eta}_i}\right] = (1-\alpha)\log P$ bits are required in the digitised reconstructed interference. After digitising the reconstructed interference, Tx-1 and Tx-2 will put the symbols into common symbols c_1 and c_2, where $\mathbb{E}[|c_1|^2] \doteq \mathbb{E}[|c_2|^2] \doteq P$.

Time Slot 2

In this time slot, Tx-1 forwards a new private symbol $s_1^{(3)}$ as well as common symbol c_2. At the same time, $s_2^{(3)}$ is sent as a private symbol from Tx-2. Overall, the transmitted symbols are given

by,

$$\mathbf{x}_1(2) = \begin{bmatrix} c_2 \\ \\ 0 \end{bmatrix} + \mathbf{q}_{21}^{\perp}(2)s_1^{(3)}$$
$$\mathbf{x}_2(2) = \mathbf{q}_{12}^{\perp}(2)s_2^{(3)},$$

with power allocation $\mathbb{E}[|s_1^{(3)}|^2] \doteq \mathbb{E}[|s_2^{(3)}|^2] \doteq P^{\alpha}$ and $\mathbb{E}[|c_2|^2] \doteq P$. Recall that c_2 contains all the information in $\bar{\tilde{\eta}}_2$. The received signals at Rx-1 and Rx-2 as well as power of each term are given by,

$$y_1(2) = \underbrace{h_{11,1}(2)c_2}_{P} + \underbrace{\mathbf{h}_{11}(2)\mathbf{q}_{21}^{\perp}(2)s_1^{(3)}}_{P^{\alpha}} + \underbrace{\mathbf{h}_{12}(2)\mathbf{q}_{12}^{\perp}(2)s_2^{(3)}}_{P^0} + n_1(2),$$
$$y_2(2) = \underbrace{h_{21,1}(2)c_2}_{P} + \underbrace{\mathbf{h}_{22}(2)\mathbf{q}_{12}^{\perp}(2)s_2^{(3)}}_{P^{\alpha}} + \underbrace{\mathbf{h}_{21}(2)\mathbf{q}_{21}^{\perp}(2)s_1^{(3)}}_{P^0} + n_2(2),$$

where $h_{ij,1}(t)$ denotes the channel from the fist antenna of Tx-j to Rx-i. With infinite SNR, for both receivers, c_2 can be decoded at the rate of $(1-\alpha)\log P$ bits/s by regarding all the other terms as noise. Consequently, all the information in c_2 can be decoded within one time slot. After subtracting the terms containing c_2, both receivers can decode their private symbols successfully because power of all the remaining terms are at noise level P^0.

Time Slot 3

In the third time slot, Tx-2 transmits c_1 as a common symbol. Meanwhile, two private symbols $s_1^{(4)}$ and $s_2^{(4)}$ are sent by Tx-1 and Tx-2, respectively. Thus, the symbols transmitted are given by,

$$\mathbf{x}_1(3) = \mathbf{q}_{21}^{\perp}(3)s_1^{(4)}$$
$$\mathbf{x}_2(3) = \begin{bmatrix} c_1 \\ \\ 0 \end{bmatrix} + \mathbf{q}_{12}^{\perp}(3)s_2^{(4)},$$

where $\mathbb{E}[|s_1^{(4)}|^2] \doteq \mathbb{E}[|s_2^{(4)}|^2] \doteq P^{\alpha}$ and $\mathbb{E}[|c_1|^2] \doteq P$. The received signals are given by

$$y_1(3) = h_{12,1}(3)c_1 + \mathbf{h}_{11}(3)\mathbf{q}_{21}^{\perp}(3)s_1^{(4)} + \mathbf{h}_{12}(3)\mathbf{q}_{12}^{\perp}(3)s_2^{(4)} + n_1(3)$$

$$y_2(3) = h_{22,1}(3)c_1 + \mathbf{h}_{22}(3)\mathbf{q}_{12}^{\perp}(3)s_2^{(4)} + \mathbf{h}_{21}(3)\mathbf{q}_{21}^{\perp}(3)s_1^{(4)} + n_2(3).$$

Similar to time slot 2, Rx-1 and Rx-2 can recover the common symbol c_1 by assuming all the other terms as noise. After subtracting the terms containing c_1, each receiver can decode their desired symbol, i.e., $s_1^{(4)}$ and $s_2^{(4)}$.

Calculate the Degrees-of-freedom

Because of the symmetry in Scheme $\mathcal{X}_1$, Rx-1 and Rx-2 achieve the same DoF. Having in mind that at the end of the third phase, Rx-1 have knowledge of all the received signals, the achievable rate at Rx-1 can be decomposed into two terms, i.e.,

$$\begin{aligned} I\left(\mathbf{s}_1, s_1^{(3)}, s_1^{(4)}; \{y_1(t)\}_{t=1}^3 | \mathfrak{H}\right) =& I\left(\mathbf{s}_1; \{y_1(t)\}_{t=1}^3 | \mathfrak{H}\right) \\ &+ I\left(s_1^{(3)}, s_1^{(4)}; \{y_1(t)\}_{t=1}^3 | \mathfrak{H}\right), \end{aligned} \tag{4.3}$$

where $I(;)$ is the mutual information and $\mathfrak{H}$ denotes all the CSI at Rx-1 as well as all the channel estimates at the transmitters. By (4.3), the total mutual information is divided into the mutual information of a 2×2 MIMO channel and the mutual information of two SISO channels.

DoF of 2×2 MIMO Channel: The MIMO channel helps to decode $\mathbf{s}_1$, which is sent in the first time slot and also included c_2. By defining $\bar{y}_1(1) = P^{\frac{\alpha}{2}} h_{12,1}(3) y_1(1) - y_1(3)$ and $\bar{y}_1(2) = y_1(2)$, the following equations can be obtained,

$$\begin{aligned} \bar{y}_1(1) = P^{\frac{\alpha}{2}} h_{12,1}(3)\mathbf{h}_{11}(1)\mathbf{x}_1(1) + P^{\frac{\alpha}{2}} h_{12,1}(3)(\tilde{\eta}_1 + \ddot{\eta}_1 + n_1(1)) \\ - \overbrace{\mathbf{h}_{11}(3)\mathbf{q}_{21}^{\perp} s_1^{(4)}}^{\mathcal{Q}_{11}} - \overbrace{\mathbf{h}_{12}(3)\mathbf{q}_{12}^{\perp}(3)s_2^{(4)}}^{\mathcal{Q}_{12}} - n_1(3) \\ \bar{y}_1(2) = P^{\frac{\alpha}{2}} h_{11,1}(2)\check{\mathbf{h}}_{21}(1)\mathbf{x}_1(1) - \overbrace{h_{11,1}(2)\tilde{\eta}_2}^{\mathcal{Q}_{21}} \\ + \overbrace{\mathbf{h}_{11}(2)\mathbf{q}_{21}^{\perp} s_1^{(3)}}^{\mathcal{Q}_{22}} + \overbrace{\mathbf{h}_{12}(2)\mathbf{q}_{12}^{\perp}(2)s_2^{(3)}}^{\mathcal{Q}_{23}} + n_1(2), \end{aligned}$$

As a result, $\bar{y}_2(1)$ and $\bar{y}_2(2)$ form an effective 2×2 MIMO channel with reduced interference,

which is given by,

$$\begin{bmatrix} \bar{y}_2(1) \\ \\ \bar{y}_2(2) \end{bmatrix} = P^{\frac{\alpha}{2}} \begin{bmatrix} h_{12,1}(3)\mathbf{h}_{11}(1) \\ \\ h_{11,1}(2)\check{\mathbf{h}}_{21}(1) \end{bmatrix} [\mathbf{q}_{21}(1) \;\; \mathbf{q}_{21}^{\perp}(1)]\mathbf{s}_1 + \begin{bmatrix} h_{12,1}P^{\frac{\alpha}{2}}(\tilde{\eta}_1 + \ddot{\eta}_1 + n_1(1)) - \mathcal{Q}_{11} - \mathcal{Q}_{12} - n_1(3) \\ \\ -\mathcal{Q}_{21} + \mathcal{Q}_{22} + \mathcal{Q}_{23} + n_1(2) \end{bmatrix},$$

where the effective channel matrix, denoted as $\mathbf{H}_e$, can be written as,

$$\mathbf{H}_e = \begin{bmatrix} h_{12,1}(3)\mathbf{h}_{11}(1) \\ \\ h_{11,1}(2)\check{\mathbf{h}}_{21}(1) \end{bmatrix} [\mathbf{q}_{21}(1) \;\; \mathbf{q}_{21}^{\perp}(1)]$$

and the effective noise vector $\mathbf{n}_e$ is given by,

$$\mathbf{n}_e = \begin{bmatrix} h_{12,1}(3)P^{\frac{\alpha}{2}}(\tilde{\eta}_1 + \ddot{\eta}_1 + n_1(1)) - \mathcal{Q}_{11} - \mathcal{Q}_{12} - n_1(3) \\ \\ -\mathcal{Q}_{21} + \mathcal{Q}_{22} + \mathcal{Q}_{23} + n_1(2) \end{bmatrix}.$$

By considering the power allocation during the scheme, components in the effective noise term have power allocation $\mathbb{E}[|\mathcal{Q}_{11}|^2] \doteq \mathbb{E}[|\mathcal{Q}_{21}|^2] \doteq \mathbb{E}[|\mathcal{Q}_{22}|^2] \doteq P^{\alpha}$, $\mathbb{E}[|\mathcal{Q}_{23}|^2] \doteq \mathbb{E}[|\mathcal{Q}_{12}|^2] \doteq P^0$, and $\mathbb{E}[|h_{12,1}P^{\frac{\alpha}{2}}(\tilde{\eta}_1 + \ddot{\eta}_1 + n_1(1))|^2] \doteq P^{1+\alpha-\beta}$. Because $1 + \alpha - \beta \geq \alpha$ in all cases, $\mathbb{E}[\mathbf{n}_e\mathbf{n}_e^H] \doteq \mathrm{diag}\{P^{1+\alpha-\beta}, P^{\alpha}\}$. The achievable rate of the equivalent MIMO channel is given by,

$$\begin{aligned} I\left(\mathbf{s}_1; \bar{y}_1(1), \bar{y}_1(2)|\mathfrak{H}\right) &= I\left(P^{\frac{\alpha}{2}}\mathbf{H}_e\mathbf{s}_1; \bar{y}_1(1), \bar{y}_1(2)|\mathfrak{H}\right) \\ &= h(P^{\frac{\alpha}{2}}\mathbf{H}_e\mathbf{s}_1|\mathfrak{H})) - h\left(\mathbf{n}_e|\bar{y}(1), \bar{y}(2), \mathfrak{H}\right) \\ &= \mathbb{E}\left[\log\det(P^{\alpha}\mathbf{N}_1^{-1}\mathbf{H}_e\mathbf{K}_e\mathbf{H}_e)\right], \end{aligned}$$

where $\mathbf{K}_e = \mathbb{E}[\mathbf{s}_1\mathbf{s}_1^H] \doteq \mathrm{diag}\{P^{1-\alpha}, P\}$, $\mathbf{N}_1 = \mathbb{E}[\mathbf{n}_e\mathbf{n}_e^H] \doteq \mathrm{diag}\{P^{1+\alpha-\beta}, P^{\alpha}\}$, and $h(\cdot)$ is the differential entropy. Thus, the mutual information is given by,

$$I\left(\mathbf{s}_1; \bar{y}_1(1), \bar{y}_1(2)|\mathfrak{H}\right) = (1 + \beta - \alpha)\log P.$$

With the definition of DoF as in (2.2), the effective MIMO channel achieves $1+\beta-\alpha$ DoF in 3 time slots.

DoF of Effective SISO Channel: The effective SISO channel helps to decode the two private symbols $s_1^{(3)}$ and $s_1^{(4)}$ transmitted from Tx-1 to Rx-1 at the second and third time slot, respectively. Firstly, the following two new symbols are defined:

$$\tilde{y}_1(2) = y_1(2) - h_{11,1}(2)c_2,$$
$$\tilde{y}_1(3) = y_1(3) - h_{12,1}(2)c_1.$$

Hence, the mutual information of the two SISO channels in (4.3) can be rewritten as $I(s_1^{(3)}, s_1^{(4)}; \tilde{y}_1(2), \tilde{y}_1(3)|\mathfrak{H})$. The corresponding mutual information can be given by,

$$I(u_3, u_4; \tilde{y}(2), \tilde{y}(3)|\mathfrak{H}) = \mathbb{E}\log\left(1 + \frac{|\mathbf{h}_{11}(2)\mathbf{q}_{21}^{\perp}(2)s_1^{(3)}|^2}{|\mathbf{h}_{12}(2)\mathbf{q}_{12}^{\perp}(2)s_2^{(3)}|^2 + |n_1(2)|^2}\right) + \mathbb{E}\log\left(1 + \frac{|\mathbf{h}_{11}(2)\mathbf{q}_{21}^{\perp}(2)s_1^{(4)}|^2}{|\mathbf{h}_{12}(2)\mathbf{q}_{12}^{\perp}(2)s_2^{(4)}|^2 + |n_1(3)|^2}\right).$$

Since,

$$\mathbb{E}\left[|\mathbf{h}_{11}(2)\mathbf{q}_{21}^{\perp}(2)s_1^{(3)}|^2\right] \doteq \mathbb{E}\left[|\mathbf{h}_{11}(2)\mathbf{q}_{21}^{\perp}(2)s_1^{(4)}|^2\right] \doteq P^{\alpha},$$
$$\mathbb{E}\left[|\mathbf{h}_{12}(2)\mathbf{q}_{12}^{\perp}(2)s_2^{(3)}|^2\right] \doteq \mathbb{E}\left[|\mathbf{h}_{12}(2)\mathbf{q}_{12}^{\perp}(2)s_2^{(4)}|^2\right] \doteq P^{0},$$

the mutual information will have the following value:

$$I\left(s_1^{(3)}, s_1^{(4)}; \tilde{y}_1(2), \tilde{y}_1(3)|\mathfrak{H}\right) = 2\alpha\log P.$$

Hence, the total DoF achieved by Rx-1 through the effective MIMO channel and the two effective SISO channels over three time slots is given by,

$$d_1 = \frac{1}{3}\lim_{P\to\infty}\frac{I\left(\mathbf{s}_1, s_1^{(3)}, s_1^{(4)}; \{y(t)\}_{t=1}^{3}|\mathfrak{H}\right)}{\log P} = \frac{1+\beta-\alpha+2\alpha}{3} = \frac{1+\alpha+\beta}{3}.$$

Symmetrically, DoF achieved by Rx-2 is $d_2 = \frac{1+\alpha+\beta}{3}$.

4.3.3 Scheme $\mathcal{X}_2$ for Vertex $A^{(1)}$

As this scheme only takes one time slot, all the time indices for symbols are ignored. Hence, the symbol vectors to be sent are given as follows:

$$\mathbf{x}_1 = \begin{bmatrix} s_1^{(1)} \\ \\ 0 \end{bmatrix} + \mathbf{q}_{21}^{\perp} s_1^{(2)}$$
$$\mathbf{x}_2 = \mathbf{q}_{12}^{\perp} s_2^{(1)},$$

where the power is allocated as $\mathbb{E}[|s_1^{(1)}|^2] \doteq P$ and $\mathbb{E}[|s_1^{(2)}|^2] \doteq E[|s_2^{(1)}|^2] \doteq P^{\alpha}$. The received signals at Rx-1 and Rx-2 are given by,

$$y_1 = \underbrace{h_{11,1} s_1^{(1)}}_{P} + \underbrace{\mathbf{h}_{11}\mathbf{q}_{21}^{\perp} s_1^{(2)}}_{P^{\alpha}} + \underbrace{\mathbf{h}_{12}\mathbf{q}_{12}^{\perp} s_2^{(1)}}_{P^0} + n_1,$$
$$y_2 = \underbrace{h_{21,1} s_1^{(1)}}_{P} + \underbrace{\mathbf{h}_{22}\mathbf{q}_{12}^{\perp} s_2^{(1)}}_{P^{\alpha}} + \underbrace{\mathbf{h}_{21}\mathbf{q}_{21}^{\perp} s_1^{(2)}}_{P^0} + n_2,$$

where $h_{ij,1}$ denotes the channel from the fist antenna of Tx-j to Rx-i and $\mathbf{q}_{ij}^{\perp}$ lies in the orthogonal space of $\hat{\boldsymbol{h}}_{ij}$. Both receivers can first decode $s_1^{(1)}$ and then decode their private symbols using SIC approach like in Scheme $\mathcal{X}_1$.

Therefore, the DoF achieved by decoding $s_1^{(1)}$ is $1-\alpha$ for both Rx-1 and Rx-2, but it is only intended for Rx-1. For both Rx-1 and Rx-2, the DoF achieved when decoding $s_1^{(2)}$ or $s_2^{(1)}$ is α. In this way, the DoF achieved by Rx-1 are $(1-\alpha+\alpha)=1$ in total and α by Rx-2. As a result, the DoF pair $(1,\alpha)$ is achieved. Instead of sending $s_1^{(1)}$ by Tx-1, Tx-2 can send private symbol for Rx-2 with power P using the first antenna. In this way, the DoF pair $(\alpha,1)$ is achieved.

Scheme $\mathcal{X}_1$ and Scheme $\mathcal{X}_2$ can be straightforwardly extended to BC. Simply letting the BS send the combination of the symbol vectors at the two transmitters in each time slot, i.e., $\mathbf{x}_1(t)+\mathbf{x}_2(t)$, will be enough.

4.4 Improved Scheme for $\mathcal{S}^{(1)}_{MAT}$ with Imperfect Current and Delayed CSIT

According to Theorem 3, although Scheme $\mathcal{X}_1$ obtains *two* DoF in total when $\alpha = \beta = 1$, it achieves lower than *one* DoF when $\alpha + \beta < \frac{1}{2}$. That is to say, Scheme $\mathcal{X}_1$ loses to the TDMA scheme when α and β are small. This is an important defect because it means that the current and delayed CSIT are harmful to the DoF region as long as their qualities are low. In order to tackle this issue, authors in [60] proposed an advanced scheme based on the Markov process to achieve higher DoF in the presence of imperfect current and delayed CSIT in a 2-user MISO BC. It will be shown that the advanced scheme achieves strictly larger than $\frac{1+\alpha+\beta}{3}$ DoF per user once $\beta < 1$ and at least *one* DoF can be achieved in total. Therefore, Scheme $\mathcal{X}_1$ can be improved.

Authors in [62] extended the MAT scheme in $\mathcal{S}^{(1)}_{MAT}$ to the time-correlated BC with imperfect current and perfect delayed CSIT. Afterwards, the authors in [64] further extended [62] to the MIMO IC where $M \geq 2N$. Similarly, the objective of this section is to extend the schemes in [60] to the IC and to consider the MIMO setting $M \geq 2N$.

4.4.1 Degrees of Freedom Region

Theorem 4 *The DoF region achieved by the MIMO IC ($M \geq 2N$) with both imperfect delayed and current CSIT when $\beta \geq \frac{1+2\alpha}{3}$ is given by,*

$$\begin{aligned} d_1 \leq N, \qquad d_2 \leq N, \\ 2d_1 + d_2 \leq N(2+\alpha), \\ d_1 + 2d_2 \leq N(2+\alpha), \end{aligned}$$

and when $\beta < \frac{1+2\alpha}{3}$,

$$\begin{aligned} d_1 \leq N, \quad d_2 \leq N, \\ (1+\beta-2\alpha)d_1 + (1-\beta)d_2 \leq N(1+\beta)(1-\alpha), \\ (1-\beta)d_1 + (1+\beta-2\alpha)d_2 \leq N(1+\beta)(1-\alpha). \end{aligned}$$

where d_i is the DoF achieved by Rx-i.

The achievable DoF region can be represented as the polygon in Figure 4.2.

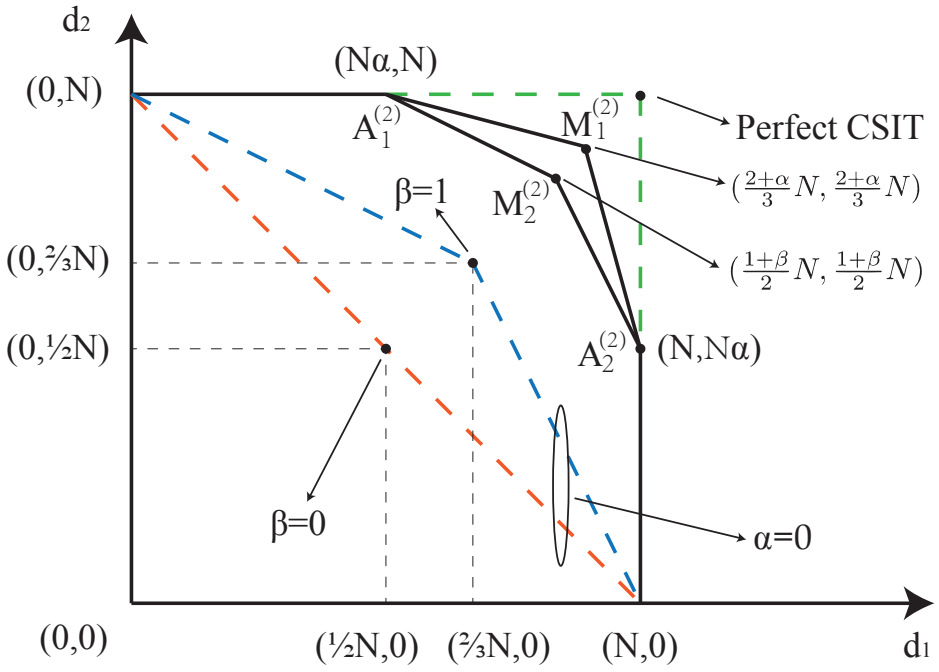


Figure 4.2: The DoF region of 2-user MIMO IC using the advanced scheme.

Remark: The point (N, N), associated with perfect CSIT, i.e., $\alpha = 1$, is achieved by the ZF scheme. The line linking the points $(N, 0)$ and $(0, N)$ can be obtained by TDMA, while the point $(\frac{N}{2}, \frac{N}{2})$ is also achieved by the proposed scheme when $\alpha = 0, \beta = 0$. When $N = 1$, the scheme reduces to what is provided in [60] and when $\alpha = 0, \beta = 1, N = 1$, the DoF region reduces to that can be achieved with the MAT scheme. When $\beta \geq \frac{1+2\alpha}{3}$, Vertex $M_1^{(2)}$ can be achieved and it outperforms Scheme $\mathcal{X}_1$ by $\frac{1-\beta}{3}$ DoF per user when $N = 1$. When $\beta < \frac{1+2\alpha}{3}$, Vertex $M_2^{(2)}$ can be achieved and it outperforms Scheme $\mathcal{X}_1$ by $\frac{1+\beta-2\alpha}{6}$ DoF per user when $N = 1$. Therefore, it is safe to draw the conclusion that Scheme $\mathcal{X}_3$ strictly outperforms Scheme $\mathcal{X}_1$ once $\beta < 1$.

One important conclusion that can be derived from Figure 4.2 is that, when β is greater than $\frac{1+2\alpha}{3}$, changing its value does not influence the DoF region, i.e., additional knowledge of the delayed CSIT is of no profit. However, when $\beta < \frac{1+2\alpha}{3}$, both α and β will affect the DoF region. Because larger bandwidth is required in the CSI feedback link to obtain higher quality of the delayed CSIT, keeping β no greater than $\frac{1+2\alpha}{3}$ is an economical solution for a practical system.

4.4.2 Scheme $\mathcal{X}_3$ for Vertices $M_1^{(2)}$ and $M_2^{(2)}$

The corresponding scheme is divided into S phases and each phase is evenly divided into 2 sub-phases. It will be shown that the highest DoF are achieved when $S \to \infty$. The number of time slots in Phase i is $2\hat{T}_i$.

Phase 1

The length of Phase 1 is $2\hat{T}_1$ and the two sub-phases have length $\hat{T}_1$, i.e., first and second sub-phases occupy time slots t where $1 \leq t \leq \hat{T}_1$ and $\hat{T}_1 + 1 \leq t \leq 2\hat{T}_1$, respectively.

Sub-phase 1**:** In this sub-phase, the private symbol vector from Tx-1 to Rx-1 is $\mathbf{s}_1(1,t) = [\mathbf{s}_1^{(1)}(1,t) \;\; \mathbf{s}_1^{(2)}(1,t)]^T$, where $\mathbf{s}_1^{(1)}(1,t), \mathbf{s}_1^{(2)}(1,t) \in \mathbb{C}^{N\times 1}$. Common symbol vector $\mathbf{c}_1(1,t) \in \mathbb{C}^{N\times 1}$ consists of the extra private symbols for Rx-1, which can be decoded by both receivers but Rx-2 regards it as interference. The private symbol vector for Rx-2 from Tx-2 is $\mathbf{s}_2(1,t) = [\mathbf{s}_2^{(1)}(1,t) \; \mathbf{s}_2^{(2)}(1,t)]^T$, where $\mathbf{s}_2^{(1)}(1,t), \mathbf{s}_2^{(2)}(1,t) \in \mathbb{C}^{N\times 1}$. The power is allocated as $\mathbb{E}[\|\mathbf{s}_1^{(1)}(1,t)\|^2] \doteq \mathbb{E}[\|\mathbf{s}_2^{(1)}(1,t)\|^2] \doteq P^{\beta-\alpha}$, $\mathbb{E}[\|\mathbf{s}_1^{(2)}(1,t)\|^2] \doteq \mathbb{E}[\|\mathbf{s}_2^{(2)}(1,t)\|^2] \doteq P^{\beta}$ and $\mathbb{E}[\|\mathbf{c}_1(1,t)\|^2] \doteq P$. The symbols are beamformed at Tx-1 and Tx-2 as,

$$\mathbf{x}_1(1,t) = \begin{bmatrix} \mathbf{c}_1(1,t) \\ \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{Q}_{21}(1,t) & \mathbf{Q}_{21}^{\perp}(1,t) \end{bmatrix} \mathbf{s}_1(1,t),$$
$$\mathbf{x}_2(1,t) = \begin{bmatrix} \mathbf{Q}_{12}(1,t) & \mathbf{Q}_{12}^{\perp}(1,t) \end{bmatrix} \mathbf{s}_2(1,t),$$

where that $\mathbf{Q}_{ij}(s,t)$ and $\mathbf{Q}_{ij}^{\perp}(s,t)$ lies in the range and orthogonal space of $\hat{\mathbf{H}}_{ij}(s,t)$, respectively. To simplify the coming descriptions, the private part of $\mathbf{x}_1(1,t)$ is defined as $\bar{\mathbf{x}}_1(1,t)$, i.e.,

$$\bar{\mathbf{x}}_1(1,t) = [\mathbf{Q}_{21}(1,t) \quad \mathbf{Q}_{21}^{\perp}(1,t)]\mathbf{s}_1(1,t).$$

The corresponding received symbols are given by,

$$\mathbf{y}_1(1,t) = \mathbf{H}_{11,1}(1,t)\mathbf{c}_1(1,t) + \mathbf{H}_{11}(1,t)\bar{\mathbf{x}}_1(1,t) + \boldsymbol{\eta}_1(1,t) + \mathbf{n}_1(1,t),$$
$$\mathbf{y}_2(1,t) = \mathbf{H}_{21,1}(1,t)\mathbf{c}_1(1,t) + \mathbf{H}_{22}(1,t)\mathbf{x}_2(1,t) + \boldsymbol{\eta}_2(1,t) + \mathbf{n}_2(1,t),$$

where $\mathbf{H}_{ij,1}(s,t) \in \mathbb{C}^{N\times N}$ is the left part of matrix $\mathbf{H}_{ij}(s,t)$ and $\boldsymbol{\eta}_i$ represents the interference

at Rx-i. Both receivers can decode $\mathbf{c}_1(1,t)$ at the rate of $N(1-\beta)\log P$ bits/s by treating other terms as noise but it is only desired by Rx-1. The interference terms can be decomposed into reconstructable part and unreconstructable parts as follows:

$$\boldsymbol{\eta}_1(1,t) = \mathbf{H}_{12}(1,t)\mathbf{x}_2(1,t) = \overbrace{\check{\mathbf{H}}_{12}^{H}(1,t)\mathbf{x}_2(1,t)}^{\check{\boldsymbol{\eta}}_1(1,t)} + \overbrace{\ddot{\mathbf{H}}_{12}(1,t)\mathbf{x}_2(1,t)}^{\ddot{\boldsymbol{\eta}}_1(1,t)},$$

and

$$\boldsymbol{\eta}_2(1,t) = \mathbf{H}_{21}(1,t)\bar{\mathbf{x}}_1(1,t) = \overbrace{\check{\mathbf{H}}_{12}(1,t)\bar{\mathbf{x}}_1(1,t)}^{\check{\boldsymbol{\eta}}_2(1,t)} + \overbrace{\ddot{\mathbf{H}}_{12}^{H}(1,t)\bar{\mathbf{x}}_1(1,t)}^{\ddot{\boldsymbol{\eta}}_2(1,t)},$$

where $\check{\boldsymbol{\eta}}_i(s,t)$ is the reconstructable part and $\ddot{\boldsymbol{\eta}}_i(s,t)$ is the corresponding reconstruction error at time slot t of Phase s. Power of the reconstructed interference $\check{\boldsymbol{\eta}}_1(1,t)$ is given by,

$$\begin{aligned}\check{\boldsymbol{\eta}}_1(1,t) =& \check{\mathbf{H}}_{12}(1,t)\mathbf{x}_2(1,t) = \check{\mathbf{H}}_{12}(1,t)[\mathbf{Q}_{12}(1,t)\ \mathbf{Q}_{12}^{\perp}(1,t)]\mathbf{s}_1(1,t)\\ =& \underbrace{\check{\mathbf{H}}_{12}(1,t)\mathbf{Q}_{12}(1,t)\mathbf{s}_1^{(1)}(1,t)}_{P^{\beta-\alpha}} + \underbrace{\check{\mathbf{H}}_{12}(1,t)\mathbf{Q}_{12}^{\perp}(1,t)\mathbf{s}_1^{(2)}(1,t)}_{P^{0}}.\end{aligned}$$

Because $\beta \geq \alpha$, the left term is the dominant one. Therefore, the reconstructed interference power is $\mathbb{E}\left[\|\check{\boldsymbol{\eta}}_1(1,t)\|^2\right] \doteq P^{\beta-\alpha}$. On the other hand, the error of the reconstructed interference is given by,

$$\begin{aligned}\ddot{\boldsymbol{\eta}}_1(1,t) =& \ddot{\mathbf{H}}_{12}(1,t)\mathbf{x}_2(1,t) = \ddot{\mathbf{H}}_{12}(1,t)[\mathbf{Q}_{12}(1,t)\ \mathbf{Q}_{12}^{\perp}(1,t)]\mathbf{s}_1(1,t)\\ =& \underbrace{\ddot{\mathbf{H}}_{12}(1,t)\mathbf{Q}_{12}(1,t)\mathbf{s}_1^{(1)}(1,t)}_{P^{-\alpha}} + \underbrace{\ddot{\mathbf{H}}_{12}(1,t)\mathbf{Q}_{12}^{\perp}(1,t)\mathbf{s}_1^{(1)}(1,t)}_{P^{0}}.\end{aligned}$$

Hence, variance of the error when reconstructing the interferences should be $\mathbb{E}\left[\|\ddot{\boldsymbol{\eta}}_1(1,t)\|^2\right] \doteq P^0$, which is comparable to the noise level. Similar analysis holds for the interference $\boldsymbol{\eta}_2(1,t)$, i.e., $\mathbb{E}\left[\|\check{\boldsymbol{\eta}}_2(1,t)\|^2\right] \doteq P^{\beta-\alpha}$ and $\mathbb{E}\left[\|\ddot{\boldsymbol{\eta}}_2(1,t)\|^2\right] \doteq P^0$.

Sub-phase 2**:** This sub-phase has similar function compared to the first sub-phase and the transmitted symbols are given by,

$$\begin{aligned}\mathbf{x}_1(1,t) &= [\mathbf{Q}_{21}(1,t) \quad \mathbf{Q}_{21}^{\perp}(1,t)]\mathbf{s}_1(1,t),\\ \mathbf{x}_2(1,t) &= \begin{bmatrix}\mathbf{c}_2(1,t)\\ \\ 0\end{bmatrix} + [\mathbf{Q}_{12}(1,t) \quad \mathbf{Q}_{12}^{\perp}(1,t)]\mathbf{s}_2(1,t),\end{aligned}$$

where $\mathbf{c}_2(1,t)$ contains private symbols intended for Rx-2. At the end of each time slot of Sub-phase 2, interference terms are reconstructed in the same way as in the first sub-phase.

Digitising Interference: At the end of the first phase, transmitters have reconstructed interference with the obtained delayed CSIT. Then, they are able to digitalise and encode the reconstructed interference to make room for new private symbols. The digitisation process and the accompanying digitisation error are given by,

$$\check{\boldsymbol{\eta}}_i(1,t) = \bar{\check{\boldsymbol{\eta}}}_i(1,t) + \tilde{\boldsymbol{\eta}}_i(1,t),$$

where $\bar{\check{\boldsymbol{\eta}}}_i(s,t)$ and $\tilde{\boldsymbol{\eta}}_i(s,t)$ denote the digitised value and the digitisation error, respectively, at time slot t of Phase s. The digitisation rate is chosen, in order to drown the digitisation error in the white noise. Hence, since the power of the delayed estimate is $\mathbb{E}\left[\|\check{\boldsymbol{\eta}}_i(1,t)\|^2\right] \doteq P^{\beta-\alpha}$, $(i=1,2)$, considering the rate distortion theorem, each $\check{\boldsymbol{\eta}}_i(1,t)$ is digitised with $N(\beta-\alpha)\log P$ bits. Hence, there are $2N\hat{T}_1(\beta-\alpha)\log P$ bits of information generated for the digitised reconstructed interference observed by one receiver in Phase 1. That it to say, there are $4N\hat{T}_1(\beta-\alpha)\log P$ bits required to be transmitted to the receivers in the next phase to make the receivers have full knowledge of the reconstructed interference terms as well as an extra observation of their private symbols.

Phase s**,** $s \in \{2,3,\ldots,S-1\}$

Phase s, which lasts for $2\hat{T}_s$ time slots, is consisted of two sub-phases with $\hat{T}_s$ time slots each.

Sub-phase 1**:** In this sub-phase, all the information in the digitised reconstructed interference generated at Rx-2 in the previous phase, i.e., $\bar{\check{\boldsymbol{\eta}}}_2(s-1,t), t=1,2,\ldots,2\hat{T}_{s-1}$, is coded in all the common symbol vectors of this phase $\mathbf{c}_1(s,t), t=1,2,\ldots,2\hat{T}_{s-1}$. Defining new private symbols $\mathbf{s}_1(s,t) = [\mathbf{s}_1^{(1)}(s,t)\ \mathbf{s}_1^{(2)}(s,t)]^T$ and $\mathbf{s}_2(s,t) = [\mathbf{s}_2^{(1)}(s,t)\ \mathbf{s}_2^{(2)}(s,t)]^T$, symbol vector sent by each transmitter is given by,

$$\mathbf{x}_1(s,t) = \begin{bmatrix} \mathbf{c}_1(s,t) \\ \\ 0 \end{bmatrix} + [\mathbf{Q}_{21}(s,t)\ \mathbf{Q}_{21}^{\perp}(s,t)]\mathbf{s}_1(s,t),$$

$$\mathbf{x}_2(s,t) = [\mathbf{Q}_{12}(s,t)\ \mathbf{Q}_{12}^{\perp}(s,t)]\mathbf{s}_2(s,t),$$

where $\mathbb{E}\left[\|\mathbf{c}_1(s,t)\|^2\right] \doteq P$, $\mathbb{E}\left[\|\mathbf{s}_1^{(1)}(s,t)\|^2\right] \doteq \mathbb{E}\left[\|\mathbf{s}_2^{(1)}(s,t)\|^2\right] \doteq P^{\beta-\alpha}$ and

$\mathbb{E}\left[\|\mathbf{s}_1^{(2)}(s,t)\|^2\right] \doteq \mathbb{E}\left[\|\mathbf{s}_2^{(2)}(s,t)\|^2\right] \doteq P^{\beta}$. The received signals in time slot t of Phase s are given by,

$$\mathbf{y}_1(s,t) = \mathbf{H}_{11,1}(s,t)\mathbf{c}_1(s,t) + \mathbf{H}_{11}(s,t)\bar{\mathbf{x}}_1(s,t) + \boldsymbol{\eta}_1(s,t) + \mathbf{n}_1(s,t),$$
$$\mathbf{y}_2(s,t) = \mathbf{H}_{21,1}(s,t)\mathbf{c}_1(s,t) + \mathbf{H}_{22}(s,t)\mathbf{x}_2(s,t) + \boldsymbol{\eta}_2(s,t) + \mathbf{n}_2(s,t),$$

where $\mathbb{E}\left[\|\mathbf{H}_{11}\bar{\mathbf{x}}_1(s,t)\|^2\right] \doteq \mathbb{E}\left[\|\mathbf{H}_{22}\mathbf{x}_2(s,t)\|^2\right] \doteq P^{\beta}$ and $\mathbb{E}\left[\|\boldsymbol{\eta}_1(s,t)\|^2\right] \doteq \mathbb{E}\left[\|\boldsymbol{\eta}_2(s,t)\|^2\right] \doteq P^{\beta-\alpha}$ just like in the first phase. By regarding all the other terms as noise, symbol $\mathbf{c}_1(s,t)$ can be decoded at the rate of $N(1-\beta)\log P$ bits/s.

Sub-phase 2**:** Symmetrically, in this sub-phase, the transmitted symbols are given by,

$$\mathbf{x}_1(s,t) = [\mathbf{Q}_{21}(s,t) \quad \mathbf{Q}_{21}^{\perp}(s,t)]\mathbf{s}_1(s,t),$$
$$\mathbf{x}_2(s,t) = \begin{bmatrix} \mathbf{c}_2(s,t) \\ \\ \mathbf{0} \end{bmatrix} + [\mathbf{Q}_{12}(s,t)\ \mathbf{Q}_{12}^{\perp}(s,t)]\mathbf{s}_2(s,t),$$

where all the $\mathbf{c}_2(s,t)$ in this sub-phase together contain all the information of digitised reconstructed interference at Rx-1 in the previous phase, i.e., $\bar{\tilde{\boldsymbol{\eta}}}_1(s-1,t), t = 1,2,\ldots,2\hat{T}_{s-1}$.

Backwards Decoding: Because $\mathbb{E}\left[\|\boldsymbol{\eta}_1(s,t)\|^2\right] \doteq \mathbb{E}\left[\|\boldsymbol{\eta}_2(s,t)\|^2\right] \doteq P^{\beta-\alpha}$, at the end of Phase s, interferences are reconstructed and digitised into $4N\hat{T}_s(\beta-\alpha)\log P$ bits in total. Taking into consideration that $N(1-\beta)$ DoF are available to decode $\mathbf{c}_i(s+1,t)$ at each receiver and each time slot, to successfully obtain all the reconstructed interference generated in the previous phase, the length of $\hat{T}_s$ should satisfy the following:

$$\hat{T}_s = \frac{2(\beta-\alpha)}{1-\beta}\hat{T}_{s-1}, \tag{4.4}$$

When receivers have full knowledge of $\mathbf{c}_i(s+1,t), i = 1,2$ in Phase $s+1$, they have full knowledge of the interference reconstructed by transmitters and another observation of their private symbols in Phase s, i.e., $\bar{\tilde{\boldsymbol{\eta}}}_1(s,t)$ and $\bar{\tilde{\boldsymbol{\eta}}}_2(s,t)$. Altogether, effective MIMO channel to decode the private symbols can be constructed at the end of Phase $s+1$, where $2 \le s \le S-1$.

Receivers obtain their private symbols in Phase s, i.e., $\mathbf{s}_1(s,t)$ and $\mathbf{s}_2(s,t)$ firstly with successive decoding method. Specifically, receivers decode $\mathbf{c}_i(s,t)$ and subtract it from the received signal. For example, in first sub-phase of Phase s, the received signals left after subtracting the terms

containing the common symbol vector are given by

$$\bar{\mathbf{y}}_1(s,t) = \mathbf{H}_{11}(s,t)\bar{\mathbf{x}}_1(s,t) + \mathbf{H}_{12}(s,t)\mathbf{x}_2(s,t) + \mathbf{n}_1(s,t),$$
$$\bar{\mathbf{y}}_2(s,t) = \mathbf{H}_{21}(s,t)\bar{\mathbf{x}}_1(s,t) + \mathbf{H}_{22}(s,t)\mathbf{x}_2(s,t) + \mathbf{n}_2(s,t).$$

In the next phase, i.e., Phase $s+1$, both receivers get $\mathbf{c}_i(s+1,t)$, which contains the digitised terms of $\check{\mathbf{H}}_{12}(s,t)\mathbf{x}_2(s,t)$ and $\check{\mathbf{H}}_{21}(s,t)\bar{\mathbf{x}}_1(s,t)$ as well as their digitisation error. For Rx-1, subtracting $\check{\mathbf{H}}_{12}(s,t)\mathbf{x}_2(s,t)$ from $\bar{\mathbf{y}}_1(s,t)$ and exploiting $\check{\mathbf{H}}_{21}(s,t)\bar{\mathbf{x}}_1(s,t)$ lead to the effective $2N \times 2N$ MIMO channel shown in (4.5), where the effective channels for the desired and interference signals are defined as $\mathbf{H}_A$ and $\mathbf{H}_B$ respectively.

$$\begin{aligned}\mathbf{y}_1 = &\underbrace{\begin{bmatrix} \mathbf{H}_{11}(s,t) \\ \\ \check{\mathbf{H}}_{21}(s,t) \end{bmatrix}}_{\mathbf{H}_A} \Big[\mathbf{Q}_{21}(s,t)\ \mathbf{Q}_{21}^{\perp}(s,t)\Big]\, \mathbf{s}_1(s,t) \\ &+ \underbrace{\begin{bmatrix} \ddot{\mathbf{H}}_{12}(s,t) \\ \\ \mathbf{0} \end{bmatrix}}_{\mathbf{H}_B} \Big[\mathbf{Q}_{12}(s,t)\ \mathbf{Q}_{12}^{\perp}(s,t)\Big]\, \mathbf{s}_2(s,t) + \underbrace{\begin{bmatrix} \widetilde{\boldsymbol{\eta}}_1(s,t) + \mathbf{n}_1(s,t) \\ \\ \widetilde{\boldsymbol{\eta}}_2(s,t) \end{bmatrix}}_{\mathbf{n}_A},\end{aligned} \tag{4.5}$$

Since $\mathbb{E}\left[\|\check{\mathbf{H}}_{21}(s,t)\mathbf{Q}_{21}^{\perp}(s,t)\|_F^2\right] \doteq P^{-\alpha}$ and the power of symbols is $\mathbb{E}[\|\mathbf{s}_1^{(1)}(s,t)\|^2] \doteq P^{\beta-\alpha}$, $\mathbb{E}\left[\|\mathbf{s}_1^{(2)}(s,t)\|^2\right] \doteq P^{\beta}$ the power of the two symbols received at Rx-1 is obtained as

$$\mathbb{E}\begin{bmatrix} \|\mathbf{H}_{A,1}\mathbf{s}_1^{(1)}(s,t)\|^2 \\ \\ \|\mathbf{H}_{A,2}\mathbf{s}_1^{(2)}(s,t)\|^2 \end{bmatrix} \doteq \begin{bmatrix} P^{\beta-\alpha} \\ \\ P^{\beta} \end{bmatrix}.$$

where $\mathbf{H}_{A,1}$ and $\mathbf{H}_{A,1}$ are the left half of $\mathbf{H}_A$ and $\mathbf{H}_B$, respectively. In the meantime, the remaining interference, white noise, and digitisation error power are all comparable in value to P^0. Thus, the total achievable DoF to decode $\mathbf{s}_1(s,t), t = 1,2,\ldots,2\hat{T}_s$ are $2N\hat{T}_s(\beta+\beta-\alpha) = 2\hat{T}_s(2\beta-\alpha)$.

Phase S

The last phase lasts for $2T_S$ times slot and T_S for each sub-phase.

Sub-phase 1**:** In this sub-phase, $\bar{\tilde{\boldsymbol{\eta}}}_2(S-1,t), t=1,2,\ldots,2T_{S-1}$ is coded into $\mathbf{c}_1(S,t)$ and sent together with two new symbol vectors $\mathbf{s}_1(S,t)$ and $\mathbf{s}_2(S,t)$, i.e.,

$$\mathbf{x}_1(S,t) = \begin{bmatrix} \mathbf{c}_1(S,t) \\ \\ \mathbf{0} \end{bmatrix} + \mathbf{Q}_{21}^{\perp}(S,t)\mathbf{s}_1(S,t),$$
$$\mathbf{x}_2(S,t) = \mathbf{Q}_{12}^{\perp}(S,t)\mathbf{s}_2(S,t).$$

The new symbol vectors have power $\mathbb{E}\left[\|\mathbf{s}_1(S,t)\|^2\right] \doteq \mathbb{E}\left[\|\mathbf{s}_2(S,t)\|^2\right] \doteq P^{\alpha}$. The received signals at this phase are given by,

$$\mathbf{y}_1(S,t) = \mathbf{H}_{11,1}(S,t)\mathbf{c}_1(S,t) + \underbrace{\mathbf{H}_{11}(S,t)\mathbf{Q}_{21}^{\perp}(S,t)\mathbf{s}_1(S,t)}_{P^{\alpha}}$$
$$+ \underbrace{\mathbf{H}_{12}(S,t)\mathbf{Q}_{12}^{\perp}(S,t)\mathbf{s}_2(S,t)}_{P^{0}} + \mathbf{n}_1(S,t),$$
$$\mathbf{y}_2(S,t) = \mathbf{H}_{22,1}(S,t)\mathbf{c}_1(S,t) + \underbrace{\mathbf{H}_{22}(S,t)\mathbf{Q}_{12}^{\perp}(S,t)\mathbf{s}_2(S,t)}_{P^{\alpha}}$$
$$+ \underbrace{\mathbf{H}_{21}(S,t)\mathbf{Q}_{21}^{\perp}(S,t)\mathbf{s}_1(S,t)}_{P^{0}} + \mathbf{n}_2(S,t),$$

which support the $N(1-\alpha)$ DoF to decode $\mathbf{c}_1(S-1,t)$. After removing the terms containing $\mathbf{c}_1(S-1,t)$, the receivers can decode their private symbols at the rate of $N\alpha \log P$ bits/s because interference is submerged in the white noise.

Sub-phase 2**:** Symmetrically, in the second sub-phase, transmitters send the following symbols:

$$\mathbf{x}_1(S,t) = \mathbf{Q}_{21}^{\perp}(S,t)\mathbf{s}_1(S,t)$$
$$\mathbf{x}_2(S,t) = \begin{bmatrix} \mathbf{c}_2(S,t) \\ \\ 0 \end{bmatrix} + \mathbf{Q}_{21}^{\perp}(S,t)\mathbf{s}_2(S,t).$$

For Phase S, because there are $2N\hat{T}_{S-1}(\beta-\alpha)\log P$ bits included in $\mathbf{c}_2(S,t)$, to successfully decode all the digitised interferences in Phase $S-1$, the number of time slots in $\hat{T}_S$ is given by,

$$\hat{T}_S = \frac{2(\beta-\alpha)}{1-\alpha}\hat{T}_{S-1}. \tag{4.6}$$

Calculate the Degrees-of-freedom:

The total DoF achieved at Rx-1 can be categorised into two kinds, one is the DoF to decode $\mathbf{c}_1(1,t)$, $\mathbf{s}_1(S,t)$ and the other one is to decodes the other symbol vectors.

Because in each of the sub-phases of Phase 1, the symbols $\mathbf{c}_1(1,t)$ are decoded with $N(1-\beta)$ DoF by Rx-1. Similarly, after decoding $\mathbf{c}_i(S,t)$ in the last phase, the DoF received for decoding the remaining private symbols are αN for each time slot. Therefore, the total DoF achieved by the effective MIMO channel to decode $\mathbf{c}_i(1,t)$, $\mathbf{s}_i(S,t)$ are $N\hat{T}_1(1-\beta)+2\alpha N\hat{T}_S$. As shown in the backwards decoding of Phase s, the DoF achieved when decoding $\mathbf{s}_i(s,t)$ in Phase s are $N\hat{T}_s(2\beta-\alpha)$, where $s=1,2,\ldots,S-1$.

The achieved DoF at Rx-1 per time slot equal to summation of all DoF achieved in all phases divided by the total number of time slots, i.e.,

$$d_1=\frac{(1-\beta)N\hat{T}_1+2\alpha N\hat{T}_S+2\sum_{i=1}^{S-1}\hat{T}_iN(2\beta-\alpha)}{2\sum_{i=1}^{S}\hat{T}_i}. \tag{4.7}$$

After taking (4.4) and (4.6) into (4.7), DoF achieved by Rx-i can be expressed as,

$$\begin{aligned}d_1&=\frac{\frac{1-\beta}{2}N\hat{T}_1+\alpha N\hat{T}_1\xi^{S-2}\zeta+N(2\beta-\alpha)\hat{T}_1\sum_{i=0}^{S-2}\xi^i}{\hat{T}_1\sum_{i=0}^{S-2}\xi^i+\hat{T}_1\xi^{S-2}\zeta}\\&=N\frac{\frac{1-\beta}{2}+\alpha\xi^{S-2}\zeta+(2\beta-\alpha)\sum_{i=0}^{S-2}\xi^i}{\sum_{i=0}^{S-2}\xi^i+\xi^{S-2}\zeta}.\end{aligned}$$

where $\xi=\frac{2(\beta-\alpha)}{1-\beta}$ and $\zeta=\frac{2(\beta-\alpha)}{1-\alpha}$. Because d_1 is increasing with S, the highest DoF are obtained when $S\to\infty$ where the achievable DoF are given by,

$$\begin{aligned}d_1&=\frac{2+\alpha}{3}N, \text{when } \beta\geq\frac{1+2\alpha}{3},\\d_1&=\frac{1+\beta}{2}N, \text{when } \beta<\frac{1+2\alpha}{3}.\end{aligned}$$

Symmetrically, Rx-2 will achieve the same DoF.

4.4.3 Scheme $\mathcal{X}_4$ for Vertices $A_1^{(2)}$ and $A_2^{(2)}$

The vertices $A_1^{(2)}=(N,N\alpha)$ and $A_2^{(2)}=(N\alpha,N)$ can be achieved by simply modifying Phase S of Scheme $\mathcal{X}_3$. Note that as this scheme only has one time slot, the time indices for the symbols

are removed. The symbols transmitted from Tx-1 and Tx-2 are given by,

$$\mathbf{x}_1 = \begin{bmatrix} \mathbf{s}_1^{(1)} \\ \\ \mathbf{0} \end{bmatrix} + \mathbf{Q}_{21}^{\perp}\mathbf{s}_1^{(2)},$$
$$\mathbf{x}_2 = \mathbf{Q}_{12}^{\perp}\mathbf{s}_2^{(1)},$$

where $\mathbf{Q}_{ij}^{\perp}$ lies in the orthogonal space of $\hat{\boldsymbol{H}}_{ij}$ and power is allocated as $\mathbb{E}\left[\|\mathbf{s}_1^{(1)}\|^2\right] \doteq P$ and $\mathbb{E}\left[\|\mathbf{s}_1^{(2)}\|^2\right] = E\left[\|\mathbf{s}_2^{(1)}\|^2\right] \doteq P^{\alpha}$. The received signals in this case are given by,

$$\mathbf{y}_1 = \underbrace{\mathbf{H}_{11,1}\mathbf{s}_1^{(1)}}_{P} + \underbrace{\mathbf{H}_{11}\mathbf{Q}_{21}^{\perp}\mathbf{s}_1^{(2)}}_{P^{\alpha}} + \underbrace{\mathbf{H}_{12}\mathbf{Q}_{12}^{\perp}\mathbf{s}_2^{(1)}}_{P^{0}} + \mathbf{n}_1,$$
$$\mathbf{y}_2 = \underbrace{\mathbf{H}_{21,1}\mathbf{s}_1^{(1)}}_{P} + \underbrace{\mathbf{H}_{22}\mathbf{Q}_{12}^{\perp}\mathbf{s}_2^{(1)}}_{P^{\alpha}} + \underbrace{\mathbf{H}_{21}\mathbf{Q}_{21}^{\perp}\mathbf{s}_1^{(2)}}_{P^{0}} + \mathbf{n}_2.$$

Both receivers can decode $\mathbf{s}_1^{(1)}$ with DoF $N(1-\alpha)$, however, it is only desired by Rx-1. After that, both receivers can decode their private messages with DoF αN. After that, Rx-1 gains $N(1-\alpha+\alpha) = N$ DoF and Rx-2 gains αN DoF.

If send a new private with power P for Rx-2 instead of sending $\mathbf{s}_1^{(1)}$, the DoF pair $(N\alpha, N)$ can be achieved.

4.5 Setting $\mathcal{S}_{MAT}^{(2)}$ with Imperfect Current and Perfect Delayed CSIT

Extending the MAT scheme to the time-correlated channels has attracted a lot of attention recently. As introduced earlier, [62] achieved $\frac{2+\alpha}{3}$ DoF in a 2-user MISO BC with imperfect current CSIT and perfect delayed CSIT. The authors in [66] studied the DoF region with evolving quality of current and delayed CSIT for a 2-user MISO BC. In [59], the DoF region for a 2-user MIMO IC with imperfect current and perfect delayed CSIT was studied. It can be seen that all the work in [59, 62, 66] has remained restricted to the 2-user case. Because the MAT scheme contains schemes for various system settings, it is meaningful to extend the other schemes to the time-correlated channel. Hence, the authors in [67] proposed the DoF outer bound and an achievability scheme in $\mathcal{S}_{MAT}^{(2)}$ with imperfect current and perfect delayed CSIT. Note that their achievability scheme was not able to achieve the DoF outer bound and how to achieve it was stated as an open problem. Therefore, the objective of this section is to provide an achievability scheme to reach the

DoF outer bound proposed in [67].

In this section, the DoF region as well as the achievability schemes for the 3-user BC are studied first and then extended to the K-user BC, where the BS has no less than NK antennas and each receiver has N antennas. Note that the most important contribution of this section is to obtain the DoF region for $\mathcal{S}^{(2)}_{MAT}$, i.e., when $N = 1$. The proposed scheme achieves $\frac{NK(1-\alpha)}{1+\frac{1}{2}+\frac{1}{3}+\cdots\frac{1}{K}} + NK\alpha$ DoF which compromise between the MAT scheme and the ZF scheme with perfect current CSIT. Moreover, the scheme achieving the asymmetric DoF tuples (receivers obtain different DoF) are proposed where one of the receivers is guaranteed N DoF.

4.5.1 Degrees of Freedom Region for 3-User Case

Theorem 5 *The achievable DoF region of a 3-user MIMO BC where $M \geq 3N$ using imperfect current CSIT as well as perfect delayed CSIT is shown in Figure 4.3 with the coordinates of the important vertices listed below:*

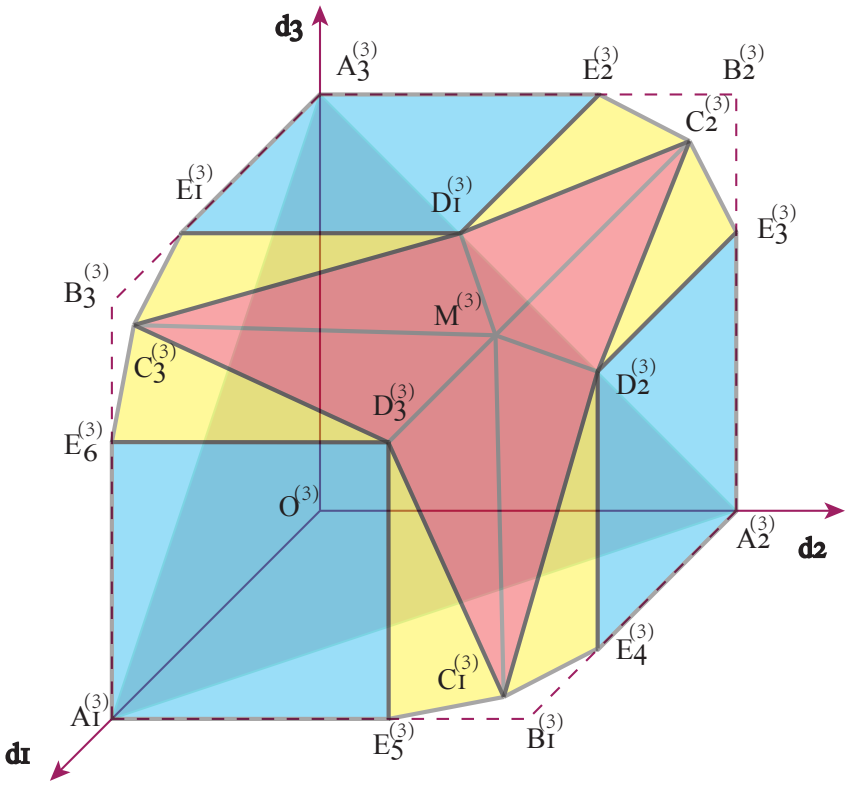


Figure 4.3: The DoF region of 3-user BC with *three* BS antennas.

$$M^{(3)} = \left(\frac{6+5\alpha}{11}N, \frac{6+5\alpha}{11}N, \frac{6+5\alpha}{11}N\right)$$

$$D_1^{(3)} = (\alpha N, \alpha N, N), D_2^{(3)} = (\alpha N, N\,\alpha N), D_3^{(3)} = (N, \alpha N, \alpha N)$$
$$C_1^{(3)} = \left(\frac{2+\alpha}{3}N, 0, \frac{2+\alpha}{3}N\right), C_2^{(3)} = \left(0, \frac{2+\alpha}{3}N, \frac{2+\alpha}{3}N\right),$$
$$C_3^{(3)} = \left(\frac{2+\alpha}{3}N, 0, \frac{2+\alpha}{3}N\right)$$

Remark: The vertices $C_1^{(3)}, C_2^{(3)}$ and $C_3^{(3)}$ are achievable using the scheme proposed in [62]. Vertices $D_1^{(3)}, D_2^{(3)}$ and $D_3^{(3)}$ achieve *three* DoF when $\alpha = 1$. Vertex $M^{(3)}$ achieves $\frac{18}{11}$ DoF when $\alpha = 0$ and *three* DoF when $\alpha = 1$, which is consistent with the MAT scheme and the ZF beamforming with perfect current CSIT.

Vertex $M^{(3)}$ with be achieved using a 3-phase achievability scheme, Scheme $\mathcal{X}_5$, where the *three* phases contain *six* time slots, *three* time slots and *two* time slots, respectively. Vertices $D_1^{(3)}, D_2^{(3)}$ and $D_3^{(3)}$ will be achieved using a one-time-slot scheme, Scheme $\mathcal{X}_6$. The achievability schemes will be discussed in detail soon.

4.5.2 Achievability Scheme $\mathcal{X}_5$ for Vertex $M^{(3)}$

Phase 1

In the first phase of Scheme $\mathcal{X}_6$, new symbol vectors for Rx-1, Rx-2 and Rx-3 are prepared as $\mathbf{s}_1$, $\mathbf{s}_2$ and $\mathbf{s}_3$, respectively, where,

$$\mathbf{s}_1 = \begin{bmatrix} \mathbf{s}_1^{(1)} \\ \mathbf{s}_1^{(2)} \\ \mathbf{s}_1^{(3)} \end{bmatrix}, \quad \mathbf{s}_2 = \begin{bmatrix} \mathbf{s}_2^{(1)} \\ \mathbf{s}_2^{(2)} \\ \mathbf{s}_2^{(3)} \end{bmatrix}, \quad \mathbf{s}_3 = \begin{bmatrix} \mathbf{s}_3^{(1)} \\ \mathbf{s}_3^{(2)} \\ \mathbf{s}_3^{(3)} \end{bmatrix}.$$

Note that all of the sub-vectors are $N \times 1$ dimensional complex vectors. Each sub-vector contains $N(1-\alpha)\log P$ bits of information and $\mathbf{E}[\|\mathbf{s}_i^{(j)}\|^2] \doteq P, \forall i, j$. When transmitting the aforementioned symbol vectors, other accompanying symbols are sent, which are denoted as $\mathbf{s}_1^{(j)} \in \mathbb{C}^{N\times 1}$, $\mathbf{s}_2^{(j)} \in \mathbb{C}^{N\times 1}$, and $\mathbf{s}_3^{(j)} \in \mathbb{C}^{N\times 1}$, where $j = 4, 5, 6$. Each of the accompanying symbol vectors contains $N\alpha \log P$ bits of information and $\mathbf{E}[\|\mathbf{s}_i^{(4)}\|^2] \doteq P^{\alpha}, \forall i$. The accompanying symbol vectors $\mathbf{s}_i^{(j)}, \forall j$ are sent at time slot $j-3$. For example, the symbol

vector sent in the first three time slots are given by,

$$\begin{aligned}
\mathbf{x}(1) &= \mathbf{s}_1 + \mathbf{Q}_{2,3}^{\perp}(1)\mathbf{s}_1^{(4)} + \mathbf{Q}_{1,3}^{\perp}(1)\mathbf{s}_2^{(4)} + \mathbf{Q}_{1,2}^{\perp}(1)\mathbf{s}_3^{(4)},\\
\mathbf{x}(2) &= \mathbf{s}_2 + \mathbf{Q}_{2,3}^{\perp}(2)\mathbf{s}_1^{(5)} + \mathbf{Q}_{1,3}^{\perp}(2)\mathbf{s}_2^{(5)} + \mathbf{Q}_{1,2}^{\perp}(2)\mathbf{s}_3^{(5)},\\
\mathbf{x}(3) &= \mathbf{s}_3 + \mathbf{Q}_{2,3}^{\perp}(3)\mathbf{s}_1^{(6)} + \mathbf{Q}_{1,3}^{\perp}(3)\mathbf{s}_2^{(6)} + \mathbf{Q}_{1,2}^{\perp}(3)\mathbf{s}_3^{(6)},
\end{aligned}$$

where $\mathbf{Q}_{i,j}^{\perp}(t) \in \mathbb{C}^{3N\times N}$ denotes the normalised beamforming matrix lying in the null space of $\mathbf{H}_i(t)$ and $\mathbf{H}_j(t)$, where $\mathbf{H}_i(t)$ denotes the channel matrix from the BS to Rx-i at time slot t. At the receiver side, the received signals in the first time slots can be expressed as,

$$\begin{aligned}
\mathbf{y}_1(1) =& \underbrace{\mathbf{H}_1(1)\mathbf{s}_1}_{P} + \underbrace{\mathbf{H}_1(1)\mathbf{Q}_{2,3}^{\perp}(1)\mathbf{s}_1^{(4)}}_{P^{\alpha}} + \underbrace{\mathbf{H}_1(1)\mathbf{Q}_{1,3}^{\perp}(1)\mathbf{s}_2^{(4)}}_{P^{0}} + \underbrace{\mathbf{H}_1(1)\mathbf{Q}_{1,2}^{\perp}(1)\mathbf{s}_3^{(4)}}_{P^{0}} + \mathbf{n}_1(1),\\
\mathbf{y}_2(1) =& \underbrace{\mathbf{H}_2(1)\mathbf{s}_1}_{P} + \underbrace{\mathbf{H}_2(1)\mathbf{Q}_{2,3}^{\perp}(1)\mathbf{s}_1^{(4)}}_{P^{0}} + \underbrace{\mathbf{H}_2(1)\mathbf{Q}_{1,3}^{\perp}(1)\mathbf{s}_2^{(4)}}_{P^{\alpha}} + \underbrace{\mathbf{H}_2(1)\mathbf{Q}_{1,2}^{\perp}(1)\mathbf{s}_3^{(4)}}_{P^{0}} + \mathbf{n}_2(1),\\
\mathbf{y}_3(1) =& \underbrace{\mathbf{H}_3(1)\mathbf{s}_1}_{P} + \underbrace{\mathbf{H}_3(1)\mathbf{Q}_{2,3}^{\perp}(1)\mathbf{s}_1^{(4)}}_{P^{0}} + \underbrace{\mathbf{H}_3(1)\mathbf{Q}_{1,3}^{\perp}(1)\mathbf{s}_2^{(4)}}_{P^{0}} + \underbrace{\mathbf{H}_3(1)\mathbf{Q}_{1,2}^{\perp}(1)\mathbf{s}_3^{(4)}}_{P^{\alpha}} + \mathbf{n}_3(1).
\end{aligned}$$

In the first time slot, Rx-i will decode $\mathbf{H}_i(1)\mathbf{s}_1$, which contains $N(1-\alpha)\log P$ bits of information, at the rate of $N(1-\alpha)\log P$ bits/s by regarding the rest of the interference as noise. Afterwards, Rx-i can remove $\mathbf{H}_i(1)\mathbf{s}_1$ from its received signal and decode their private symbol as rest of the interference submerges in noise. Using a similar method, in each time slot of Phase 1, Rx-i is able to decode $\mathbf{H}_i(t)\mathbf{s}_j, \forall j, t$ as well as their private symbols, i.e., $\mathbf{s}_i^{(j)}, j = 4, 5, 6$. In order to accomplish the following phases, the first phase is repeated once for new symbol vectors where $\bar{\mathbf{s}}_1$, $\bar{\mathbf{s}}_2$ and $\bar{\mathbf{s}}_3$ replace $\mathbf{s}_1$, $\mathbf{s}_2$ and $\mathbf{s}_3$, respectively, and $\mathbf{s}_i^{(j)}, j = 7, 8, 9$ replace $\mathbf{s}_i^{(j)}, j = 4, 5, 6$, respectively.

At the end of Phase 1, the BS has obtained perfect CSIT for the first phase, with which it can perfectly reconstruct all the interference terms. Similar to the MAT scheme, the order-2 symbols will be reconstructed using the reconstructed interference. Specifically, $\mathbf{H}_2(1)\mathbf{s}_1 + \mathbf{H}_1(2)\mathbf{s}_2$ is digitised and coded into $\mathbf{c}_{12}$, $\mathbf{H}_3(1)\mathbf{s}_2 + \mathbf{H}_1(3)\mathbf{s}_3$ is digitised and coded into $\mathbf{c}_{13}$, $\mathbf{H}_3(2)\mathbf{s}_2 + \mathbf{H}_2(3)\mathbf{s}_3$ is digitised and coded into $\mathbf{c}_{23}$. Symmetrically, $\bar{\mathbf{c}}_{13}$, $\bar{\mathbf{c}}_{23}$ and $\bar{\mathbf{c}}_{12}$ are also generated. Because each interference term $\mathbf{H}_i(t)\mathbf{s}_j$ is decoded at the rate of $N(1-\alpha)\log P$ bits/s of information, each $\mathbf{c}_{ij}$ or $\bar{\mathbf{c}}_{ij}$ also contains $N(1-\alpha)\log P$ bits. In the rest of this section $\mathbf{c}_{ij}$ and $\bar{\mathbf{c}}_{ij}$ are called order-2 common symbol vectors.

Phase 2

The symbol vectors sent in each of the three time slots of the second phase are given by,

$$\begin{aligned}
\mathbf{x}(7) &= [\mathbf{c}_{12} \;\; \bar{\mathbf{c}}_{12} \;\; 0]^T + \mathbf{Q}_{2,3}^{\perp}(7)\mathbf{s}_1^{(10)} + \mathbf{Q}_{1,3}^{\perp}(7)\mathbf{s}_2^{(10)} + \mathbf{Q}_{1,2}^{\perp}(7)\mathbf{s}_3^{(10)}, \\
\mathbf{x}(8) &= [\mathbf{c}_{13} \;\; \bar{\mathbf{c}}_{13} \;\; 0]^T + \mathbf{Q}_{2,3}^{\perp}(8)\mathbf{s}_1^{(11)} + \mathbf{Q}_{1,3}^{\perp}(8)\mathbf{s}_2^{(11)} + \mathbf{Q}_{1,2}^{\perp}(8)\mathbf{s}_3^{(11)}, \\
\mathbf{x}(9) &= [\mathbf{c}_{23} \;\; \bar{\mathbf{c}}_{23} \;\; 0]^T + \mathbf{Q}_{2,3}^{\perp}(9)\mathbf{s}_1^{(12)} + \mathbf{Q}_{1,3}^{\perp}(9)\mathbf{s}_2^{(12)} + \mathbf{Q}_{1,2}^{\perp}(9)\mathbf{s}_3^{(12)},
\end{aligned}$$

where each $\mathbf{c}_{ij}$ and $\bar{\mathbf{c}}_{ij}$ has power $\frac{P}{2}$ in order to satisfy the power constraint. For DoF analysis, only the power exponents of the received terms are cared about, which means the coefficient $\frac{1}{2}$ in power of order-2 common symbol vectors can be safely removed for the rest of the derivations.

During the three time slots, the receivers will obtain combination of $\mathbf{c}_i$ and $\bar{\mathbf{c}}_i$ with power P, desired private symbols with power P^{α} and interference with power P^0. For example, at time slot 7, Rx-1 will receive the following terms:

$$y_1(7) = \underbrace{\mathbf{H}_1(7)\begin{bmatrix} \mathbf{c}_{12} \\ \bar{\mathbf{c}}_{12} \\ 0 \end{bmatrix}}_{P} + \underbrace{\mathbf{H}_1(7)\mathbf{Q}_{2,3}^{\perp}(7)\mathbf{s}_1^{(10)}}_{P^{\alpha}} + \underbrace{\mathbf{H}_1(7)\mathbf{Q}_{1,3}^{\perp}(7)\mathbf{s}_2^{(10)}}_{P^0} + \underbrace{\mathbf{H}_1(7)\mathbf{Q}_{1,2}^{\perp}(7)\mathbf{s}_3^{(10)}}_{P^0} + \mathbf{n}_1(7).$$

As the combination of $\mathbf{c}_{12}$ and $\bar{\mathbf{c}}_{12}$ contains $N(1-\alpha)\log P$ bits of information, it can be successfully decoded by regarding rest of the terms as noise. Afterwards, the private symbol $\mathbf{s}_1^{(10)}$ can be decoded at the rate of $N\alpha \log P$ using successive decoding method. Similarly, at time slot $t, t = 7, 8, 9$, Rx-i will be able to decode the combination of $\mathbf{c}_i$ and $\bar{\mathbf{c}}_i$ as well as their private symbol vectors $\mathbf{s}_i^{(t+3)}$.

At the end of Phase 2, the BS obtains perfect CSI in the second phase. Similar to the MAT scheme, it will reconstruct the overheard interference and generate the order-3 common symbol vectors, i.e., $\mathbf{c}_{123}^{(1)}$ is coded from $\mathbf{H}_3(7)\,[\mathbf{c}_{12} \;\; \bar{\mathbf{c}}_{12} \;\; 0]^T$, $\mathbf{c}_{123}^{(2)}$ is coded from $\mathbf{H}_2(8)\,[\mathbf{c}_{13} \;\; \bar{\mathbf{c}}_{13} \;\; 0]^T$ and $\mathbf{c}_{123}^{(3)}$ is coded from $\mathbf{H}_1(9)\,[\mathbf{c}_{23} \;\; \bar{\mathbf{c}}_{23} \;\; 0]^T$. It is worth mentioning that as each combination of $\mathbf{c}_{ij}$ and $\bar{\mathbf{c}}_{ij}$ is decoded at the rate of $N(1-\alpha)\log P$ bits/s, each $\mathbf{c}_{123}^{(i)}$ contains $N(1-\alpha)\log P$ bits of information.

Phase 3

Similar to the MAT scheme, in the two time slots of Phase 3, the BS will send two linear combinations of $\mathbf{c}_{123}^{(1)}$, $\mathbf{c}_{123}^{(2)}$ and $\mathbf{c}_{123}^{(3)}$. Specifically, the symbol vector sent at time slot t, where $t = 10, 11$, is given by,

$$\mathbf{x}(t) = \begin{bmatrix} u_t\mathbf{c}_{123}^{(1)} + v_t\mathbf{c}_{123}^{(2)} + w_t\mathbf{c}_{123}^{(3)} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} + \mathbf{Q}_{2,3}^{\perp}(t)\mathbf{s}_1^{(t+3)} + \mathbf{Q}_{1,3}^{\perp}(t)\mathbf{s}_2^{(t+3)} + \mathbf{Q}_{1,2}^{\perp}(t)\mathbf{s}_3^{(t+3)}. \quad (4.8)$$

The signal received by Rx-i at time slot t is given by,

$$\begin{aligned}\mathbf{y}_i(t) =& \mathbf{H}_i(t) \begin{bmatrix} u_t\mathbf{c}_{123}^{(1)} + v_t\mathbf{c}_{123}^{(2)} + w_t\mathbf{c}_{123}^{(3)} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} + \mathbf{H}_i(t)\mathbf{Q}_{2,3}^{\perp}(t)\mathbf{s}_1^{(t+3)} \\ &+ \mathbf{H}_i(t)\mathbf{Q}_{1,3}^{\perp}(t)\mathbf{s}_2^{(t+3)} + \mathbf{H}_i(t)\mathbf{Q}_{1,2}^{\perp}(t)\mathbf{s}_3^{(t+3)} + \mathbf{n}_i(t),\end{aligned}$$

where $u_t\mathbf{c}_{123}^{(1)} + v_t\mathbf{c}_{123}^{(2)} + w_t\mathbf{c}_{123}^{(3)}$ can be decoded at the rate of $N(1-\alpha)\log P$ bits/s by regarding rest of the terms as noise. Similar to Phase 1 and Phase 2, their private symbol vectors can be decoded at the rate of $N\alpha\log P$ bits/s using the SIC approach.

Backwards Decoding and DoF Calculation

The symbols contained in $\mathbf{s}_i$ and $\bar{\mathbf{s}}_i$ can be decoded by Rx-i using the approach provided in the MAT scheme as all of the required order-2 and order-3 common symbol vectors have already been obtained. Meanwhile, at Rx-i, all the new private symbols included in $\mathbf{s}_i^{(j)}$ and $\bar{\mathbf{s}}_i^{(j)}$, where $j = 7, 8, \ldots, 14$, can be decoded in each time slot using SIC method.

On the one hand, this scheme sends $\mathbf{s}_1$, $\bar{\mathbf{s}}_1$, $\mathbf{s}_2$, $\bar{\mathbf{s}}_2$, $\mathbf{s}_3$ and $\bar{\mathbf{s}}_3$ to receivers in 11 time slots, which contains $18N(1-\alpha)\log P$ bits of information in total. Therefore, the total DoF gained for decoding $\mathbf{s}_1$, $\bar{\mathbf{s}}_1$, $\mathbf{s}_2$, $\bar{\mathbf{s}}_2$, $\mathbf{s}_3$ and $\bar{\mathbf{s}}_3$ are $\frac{18(1-\alpha)}{11}N$. On the other hand, in each of the time slots, three private symbol vectors are sent for all of the receivers and each contains $N\alpha\log P$ bits of

information. That is to say, the DoF per time slot for decoding $\mathbf{s}_i^{(j)}, j = 7, 8 \ldots, 14$ are $3N\alpha$. Summing up, the total DoF achieved by this scheme are

$$N\frac{18(1-\alpha)}{11} + 3N\alpha = N\frac{18+15\alpha}{11}.$$

Because the scheme is symmetric for Rx-1, Rx-2 and Rx-3, each receiver obtains $N\frac{6+5\alpha}{11}$ DoF.

Remark: Scheme $\mathcal{X}_5$ can not be extended to IC. This is because in the order-3 common symbol, special linear combinations of the private symbols are required. For example, $\mathbf{c}_{123}^{(1)}$ is coded based on $\mathbf{H}_3(7)\left[\mathbf{c}_{12} \;\; \bar{\mathbf{c}}_{12} \;\; 0\right]^T$, where $\mathbf{c}_{12}$ is further coded based on $\mathbf{H}_2(1)\mathbf{s}_1 + \mathbf{H}_1(2)\mathbf{s}_2$. Hence, when transmitting $\mathbf{c}_{123}^{(1)}$, it must be confirmed that the included terms $\mathbf{H}_2(1)\mathbf{s}_1$ and $\mathbf{H}_1(2)\mathbf{s}_2$ are multiplied by the same channel matrix when $\mathbf{c}_{123}^{(1)}$ arrives at each receiver. However, this is an impossible job because $\mathbf{s}_1$ and $\mathbf{s}_2$ are only known by Tx-1 and Tx-2, respectively, and the channels from the two transmitters are mutually independent.

4.5.3 Achievability Scheme $\mathcal{X}_6$ for Vertices $D_1^{(3)}$, $D_2^{(3)}$ and $D_3^{(3)}$

The Scheme $\mathcal{X}_6$ will provide a specific scheme to achieve vertex $D_1^{(3)} = (N, N\alpha, N\alpha)$ in the DoF region. This achievability scheme will only occupy one time slot, and the symbol vector that is to be sent is given by,

$$\mathbf{x} = \begin{bmatrix} \mathbf{s}_1^{(1)} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} + \mathbf{Q}_{2,3}^{\perp}\mathbf{s}_1^{(2)} + \mathbf{Q}_{1,3}^{\perp}\mathbf{s}_2^{(1)} + \mathbf{Q}_{1,2}^{\perp}\mathbf{s}_3^{(1)},$$

where $\mathbf{s}_1^{(1)} \in \mathbb{C}^{N\times 1}$ is the private symbol for Rx-1 and $\mathbb{E}[\|\mathbf{s}_1^{(1)}\|^2] \doteq P$. Symbols $\mathbf{s}_1^{(2)}$, $\mathbf{s}_2^{(1)}$ and $\mathbf{s}_3^{(1)}$, all with dimension $\mathbb{C}^{N\times 1}$, are the private symbol vectors for Rx-1, Rx-2 and Rx-3, respectively, and each has power P^{α}. The received signal at Rx-1 can be expressed as,

$$\mathbf{y}_1 = \underbrace{\mathbf{H}_{1,1}\mathbf{s}_1^{(1)}}_{P} + \underbrace{\mathbf{H}_1\mathbf{Q}_{2,3}^{\perp}\mathbf{s}_1^{(2)}}_{P^{\alpha}} + \underbrace{\mathbf{H}_1\mathbf{Q}_{1,3}^{\perp}\mathbf{s}_2^{(1)}}_{P^0} + \underbrace{\mathbf{H}_1\mathbf{Q}_{1,2}^{\perp}\mathbf{s}_3^{(1)}}_{P^0} + \mathbf{n}_1,$$

where $\mathbf{H}_{1,1}$ denotes the left $N \times N$ matrix in $\mathbf{H}_1$. Rx-1 will decode $\mathbf{s}_1^{(1)}$ at rate $N(1-\alpha)\log P$ bits/s by regarding the rest of the terms as noise. Then it can remove $\mathbf{H}_{1,1}\mathbf{s}_1^{(1)}$ and decode the other

private symbol in $\mathbf{s}_1^{(2)}$ at the rate of $N\alpha \log P$ bits/s because the other interference submerges in noise. Therefore, Rx-1 gets $N(1-\alpha)+N\alpha = N$ DoF in total.

On the other side, Rx-2 and Rx-3 will also be able to decode $\mathbf{H}_{i,1}\mathbf{s}_1^{(1)}, i = 2,3$ at rate $N(1-\alpha)\log P$ bits/s. However, they regard it as an undesired symbol vector and remove it from the received signal. Afterwards, they can decode their private symbol at the rate of $N\alpha \log P$ bits/s. Hence, each obtains $N\alpha$ DoF. As a result, the DoF vertex $(N, N\alpha, N\alpha)$ is achieved.

By changing $\mathbf{s}_1^{(1)}$ into symbol vectors for Rx-2 or Rx-3, the vertices $D_2^{(3)}$ and $D_1^{(3)}$ can be achieved, respectively.

4.5.4 Extension to the K-User Case

For the K-user setting, two achievability schemes are considered, i.e., Scheme $\mathcal{X}_7$ and Scheme $\mathcal{X}_8$. To be more specific, Scheme $\mathcal{X}_7$ will achieve the symmetric DoF vertex, i.e., all of the users get the same DoF. Scheme $\mathcal{X}_8$ will achieve the DoF vertex when one of the receivers obtains N DoF.

Achievability Scheme $\mathcal{X}_7$

Theorem 6 *In a K-user MIMO BC where the BS is equipped with KN antennas and receivers are equipped with N antennas, if the current CSIT is imperfect while the delayed CIST is perfect, the total achievable DoF are given by,*

$$\frac{NK(1-\alpha)}{1+\frac{1}{2}+\frac{1}{3}+\cdots\frac{1}{K}} + NK\alpha,$$

which are evenly allocated to the K receivers.

Remark: This achieved DoF are consistent with two well known result when $N = 1$. Specifically, when $\alpha = 0$, $\frac{1}{1+\frac{1}{2}+\cdots+\frac{1}{K}}$ DoF are achieved which can also by achieved using the MAT scheme. When $\alpha = 1$, the DoF per receiver become *one*, which can be simply achieved using the ZF beamforming method.

The MAT scheme is modified in order to achieve this DoF vertex. Specifically, Scheme $\mathcal{X}_7$ can be divided into the MAT part and the ZF part:

- In the MAT part, the MAT scheme is performed but instead of constructing each symbol

vector with $N \log P$ bits of information, it is constructed with $N(1-\alpha)\log P$ bits of information. Consequently, each order-i symbol vector used in the MAT scheme, where $i = 1, 2, \ldots, K$ will contain $N(1-\alpha)\log P$ bits of information. Using the backwards decoding approach as stated in [8], $\frac{NK(1-\alpha)}{1+\frac{1}{2}+\frac{1}{3}+\cdots\frac{1}{K}}$ DoF are achieved.

- In the ZF part, because all of the symbol vectors in the MAT part contains only $N(1-\alpha)\log P$ bits of information, they can be sent with power P, by which there will be extra space to send new private symbol vectors at each time slot. Because the BS is equipped with KN antennas, it can provide bases for K independent N-dimensional spaces. Specifically, if $\mathbf{s}_i \in \mathbb{C}^{N\times 1}$ is desired by Rx-i, the symbol vector in the ZF part is beamformed as,

$$\sum_{i=1}^{K} \mathbf{Q}^{\perp}_{1,2,\ldots,i-1,i+1\ldots,K}\mathbf{s}_i,$$

where $\mathbf{Q}^{\perp}_{1,2,\ldots,i-1,i+1\ldots,K}$ denotes the beamforming matrix in the null space of the estimated current channel matrices to all the receivers except Rx-i. If $\mathbb{E}[\|\mathbf{s}_i\|^2] \doteq P^{\alpha}$, all the overheard interference will be at noise level. Therefore, at each time slot, there are K symbol vectors containing $KN\alpha\log P$ bits of information for each of the receivers and $NK\alpha$ DoF are achieved.

Summing up, the total DoF achieved can be written as,

$$\frac{NK(1-\alpha)}{1+\frac{1}{2}+\frac{1}{3}+\cdots\frac{1}{K}} + NK\alpha,$$

which are evenly allocated to the K receivers.

Similar to Scheme $\mathcal{X}_5$, Scheme $\mathcal{X}_7$ can not be extended to IC because special combinations of private symbols are required.

Achievability Scheme $\mathcal{X}_8$

Scheme $\mathcal{X}_8$ is a straightforward extension of Scheme $\mathcal{X}_6$. Specifically, because the BS is equipped with KN antennas, it can provides K linear independent N-dimensional space and the beamforming matrices for the desired symbol vector for Rx-i are denoted as $\mathbf{Q}^{\perp}_{1,2,\ldots,i-1,i+1\ldots,K}$, where $\mathbf{Q}^{\perp}_{1,2,\ldots,i-1,i+1\ldots,K}$ is orthogonal to $\mathrm{span}(\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_{i-1}, \mathbf{H}_{i+1}, \ldots, \mathbf{H}_K)$. If the defined new symbol vectors $\mathbf{s}_i^{(1)} \in \mathbb{C}^{N\times 1}$ and $\mathbf{s}_i^{(2)} \in \mathbb{C}^{N\times 1}$ are intended for Rx-i, the symbol

vector to be sent can be written as,

$$\mathbf{x} = \mathbf{s}_i^{(1)} + \sum_{j=1}^{K} \mathbf{Q}_{1,2,\ldots,j-1,j+1\ldots,K}^{\perp} \mathbf{s}_j^{(2)},$$

where $\mathbb{E}\left[\|\mathbf{s}_i^{(1)}\|^2\right] \doteq P$ and $\mathbb{E}\left[\|\mathbf{s}_i^{(2)}\|^2\right] \doteq P^{\alpha}, \forall i$. For Rx-$k$, $k = 1, 2, \ldots, K$, the received signals are given by,

$$\mathbf{y}_k = \underbrace{\mathbf{H}_{k,1}\mathbf{s}_i^{(1)}}_{P} + \underbrace{\mathbf{H}_k \mathbf{Q}_{1,2,\ldots,k-1,k+1\ldots,K}^{\perp} \mathbf{s}_k^{(2)}}_{P^{\alpha}} + \underbrace{\mathbf{H}_k \sum_{j=1,j\neq k}^{K} \mathbf{Q}_{1,2,\ldots,j-1,j+1\ldots,K}^{\perp} \mathbf{s}_k^{(2)}}_{P^{0}} + \mathbf{n}_k.$$

Rx-k can first decode $\mathbf{H}_{k,1}\mathbf{s}_i^{(1)}$ at rate $N(1-\alpha)\log P$ bits/s by regarding the rest of the terms as noise. Then it can remove it from $\mathbf{y}_k$ and decode its own private symbol at rate $N\alpha \log P$ because the remaining interference is at noise level. Because $\mathbf{s}_i^{(1)}$ is intended for Rx-i, then the achieved DoF at Rx-i are $N(1-\alpha) + N\alpha = N$. At all the other receivers, the achieved DoF are $N\alpha$.

4.6 Setting $\mathcal{S}_{MAT}^{(3)}$ with Imperfect Current and Perfect Delayed CSIT

Until now, the MAT scheme in $\mathcal{S}_{MAT}^{(1)}$ and $\mathcal{S}_{MAT}^{(2)}$ has been extended to the time-correlated channels by [59, 62, 66] and the previous section. Hence, the objective of this section is to extend the MAT scheme in the only remaining setting, i.e., $\mathcal{S}_{MAT}^{(3)}$, to a time-correlated channel. Note that $\mathcal{S}_{MAT}^{(3)}$ is quite a special setting because the number of BS antennas is less than the number of users. In this way, the MAT scheme for $\mathcal{S}_{MAT}^{(3)}$ is quite different to the MAT scheme in $\mathcal{S}_{MAT}^{(1)}$ and $\mathcal{S}_{MAT}^{(2)}$, which leads to the differences in the ZF beamforming and the overheard interference digitisation when being extended to a time-correlated channel.

In this section, 3-user BC is considered using imperfect current CSIT and perfect delayed CSIT. It is assume that the BS is equipped with $2N$ antennas and each receiver is equipped with N antennas. The most important contribution for this section is to achieve $\frac{3+\alpha}{6}$ DoF at each receiver in $\mathcal{S}_{MAT}^{(3)}$ although the MIMO setting is considered. The DoF are consistent with the MAT scheme and the ZF beamforming scheme with perfect current CSIT. Thus, this scheme is optimal in the infinite SNR case. Similar to Section 4.5, another contribution of this section is to visualise the DoF region for a 3-user case, which clearly shows how the achievable DoF compromise among the three receivers. Different from the 2-dimensional DoF regions, its convexity can not been straightforwardly observed through the coordinates of the vertices and therefore, proof of

convexity is provided. Finally, it will be shown that the proposed scheme in this section cannot be extended to IC.

4.6.1 Degrees of Freedom Region

Theorem 7 *The DoF region achieved by the 3-user MIMO BC with perfect delayed CSIT as well as imperfect current CSIT when the BS is equipped with $2N$ antennas and the receivers are equipped with N antenna can be formed by the following inequalities:*

$$
\begin{aligned}
&d_1 \leq N,\ d_2 \leq N,\ d_3 \leq N,\\
&2d_1 + d_2 + d_3 \leq N(2+\alpha),\\
&d_1 + 2d_2 + d_3 \leq N(2+\alpha),\\
&d_1 + d_2 + 2d_3 \leq N(2+\alpha),\\
&2d_1 + 3(1+\alpha)d_2 + 3(1+\alpha)d_3 \leq N(2+\alpha)(3+\alpha),\\
&3(1+\alpha)d_1 + 2d_2 + 3(1+\alpha)d_3 \leq N(2+\alpha)(3+\alpha),\\
&3(1+\alpha)d_1 + 3(1+\alpha)d_2 + 2d_3 \leq N(2+\alpha)(3+\alpha),\\
&d_1 + d_2 + d_3 \leq 2N.
\end{aligned}
$$

The polygon formed by the planes is shown in Figure 4.4 and the important vertices are listed as follows:

$$
\begin{aligned}
&C_1^{(4)} = \left(N\frac{2+\alpha}{3}, N\frac{2+\alpha}{3}, 0\right), C_2^{(4)} = \left(0, N\frac{2+\alpha}{3}, N\frac{2+\alpha}{3}\right),\\
&C_3^{(4)} = \left(N\frac{2+\alpha}{3}, 0, N\frac{2+\alpha}{3}\right),\\
&D_1^{(4)} = (N\alpha, 0, N), D_2^{(4)} = (0, N\alpha, N),\\
&E_1^{(4)} = (0, N, N\alpha), E_2^{(4)} = (N\alpha, N, 0),\\
&F_1^{(4)} = (N, 0, N\alpha), F_2^{(4)} = (N, N\alpha, 0),\\
&M^{(4)} = \left(N\frac{3+\alpha}{6}, N\frac{3+\alpha}{6}, N\frac{3+\alpha}{6}\right).
\end{aligned}
$$

Remark: Lines $D_1^{(4)}D_2^{(4)}$, $E_1^{(4)}E_2^{(4)}$ and $F_1^{(4)}F_2^{(4)}$ are parallel to lines $C_2^{(4)}C_3^{(4)}$, $C_1^{(4)}C_2^{(4)}$ and $C_1^{(4)}C_3^{(4)}$, respectively. The DoF surface represented by triangle $A_1^{(4)}A_2^{(4)}A_3^{(4)}$ can be obtained by the TDMA scheme where N DoF can be achieved totally. All the remaining vertices except vertex

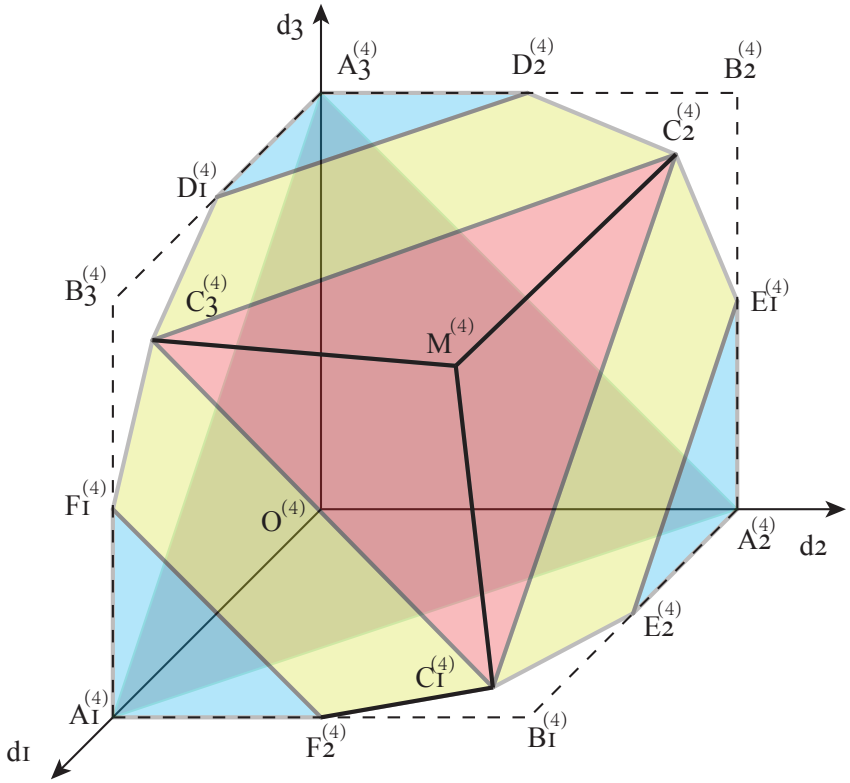


Figure 4.4: The DoF region of 3-user MIMO BC with *two* BS antennas.

$O^{(4)}$ achieve $2N$ total DoF when $\alpha = 1$, which is the achievable DoF using the ZF method with perfect current CSIT. The total DoF achieved at vertex $M^{(4)}$ is $N\frac{3+\alpha}{2}$, which is consistent with two well known results, i.e., when $\alpha = 1$, $2N$ DoF are achieved and when $\alpha = 0$, $\frac{3N}{2}$ DoF are achieved which is proven in [8]. The panels $O^{(4)}A_2^{(4)}E_1^{(4)}C_2^{(4)}D_2^{(4)}A_3^{(4)}$, $O^{(4)}A_2^{(4)}E_2^{(4)}C_1^{(4)}F_2^{(4)}A_1^{(4)}$ and $O^{(4)}A_1^{(4)}F_1^{(4)}C_3^{(4)}D_1^{(4)}A_3^{(4)}$ can be achieved using the scheme proposed in [62]. Unlike Figure 4.4, if one of the receivers is guaranteed N DoF, the other two receivers can achieve $N\alpha$ DoF in total.

Proof of Convexity

As the DoF tuple is defined as the pre-log factor of the highest achievable sum rate which is shown in (2.3), one important property of DoF regions is convexity. Therefore, the 3-dimensional DoF region will be proven to be convex. In order to proof its convexity, the following three equivalent hypotheses are proven.

- Vertex $M^{(4)}$ is above the imaginary panel $C_1^{(4)}C_2^{(4)}C_3^{(4)}$.

Proof: The panel $C_1^{(4)}C_2^{(4)}C_3^{(4)}$ can be described by the following equation

$$f_1(d_1, d_2, d_3) = 3d_1 + 3d_2 + 3d_3 - 2N(2+\alpha) = 0.$$

Take the coordination of vertex $M^{(4)}$ into $f_1(d_1, d_2, d_3)$, the following inequality can be obtained,

$$f_1\left(\frac{3+\alpha}{6}, \frac{3+\alpha}{6}, \frac{3+\alpha}{6}\right) = \frac{N}{2} - \frac{N\alpha}{2} \geq 0,$$

which means vertex $M^{(4)}$ is above the panel $C_1C_2C_3$, i.e., the polygon is convex at vertex $M^{(4)}$.

- Any point in line $C_2^{(4)}C_3^{(4)}$ is above the imaginary panel $M^{(4)}D_1^{(4)}D_2^{(4)}$.

  Proof: The panel $M^{(4)}D_1^{(4)}D_2^{(4)}$ can be described by the following equation

  $$f_2(d_1, d_2, d_3) = (\alpha-3)d_1 + (\alpha-3)d_2 + (4\alpha-6)d_3 - N(\alpha+3)(\alpha-2) = 0.$$

  The functions of line $C_2^{(4)}C_3^{(4)}$ are given by

  $$\begin{aligned} d_3 &= N\frac{2+\alpha}{3}, \\ d_1 + d_2 &= N\frac{2+\alpha}{3}. \end{aligned}$$

  Taking line $C_2^{(4)}C_3^{(4)}$ into $f_2(d_1, d_2, d_3)$, the following inequality can be obtained:

  $$f_2\left(d_1, d_2, d_3 \middle| d_3 = N\frac{2+\alpha}{3}; d_1 + d_2 = N\frac{2+\alpha}{3}\right) = 4\alpha^2 N \geq 0,$$

  which verifies this hypothesis.

- Any point in line $D_1^{(4)}D_2^{(4)}$ is above the imaginary panel $A^{(4)}C_1^{(4)}C_2^{(4)}$.

  Proof: The panel $A^{(4)}C_1^{(4)}C_2^{(4)}$ can be represented by the following equation

  $$f_3(d_1, d_2, d_3) = (1-\alpha)d_1 + (1-\alpha)d_2 + (2+\alpha)d_3 - N(2+\alpha) = 0,$$

  and the functions of line $D_1^{(4)}D_2^{(4)}$ are given by

  $$\begin{aligned} d_3 &= N, \\ d_1 + d_2 &= N\alpha. \end{aligned}$$

By taking the line function of $D_1^{(4)}D_2^{(4)}$ into the panel function, it can be observed that

$$f_3(d_1, d_2, d_3 \big| d_3 = N; d_1 + d_2 = N\alpha) = \alpha(1-\alpha)N \geq 0,$$

which also verifies the hypothesis.

With the above three hypotheses being true, the panels $A_3^{(4)}D_1^{(4)}D_2^{(4)}$, $D_1^{(4)}D_2^{(4)}C_2^{(4)}C_3^{(4)}$ and $M^{(4)}C_2^{(4)}C_3^{(4)}$ form convex polygon. Based on this result, it can be easily proven that the whole polygon is convex because of symmetry.

4.6.2 Scheme $\mathcal{X}_9$ for Vertex $M^{(4)}$

The total scheme is divided into three phases where the first two phases consist of 3 time slots and the last phase contains 2 time slots. Two symbol vectors sent to each of the receivers are $\mathbf{s}_1 = [\mathbf{s}_1^{(1)}\ \mathbf{s}_1^{(2)}]^T, \bar{\mathbf{s}}_1 = [\bar{\mathbf{s}}_1^{(1)}\ \bar{\mathbf{s}}_1^{(2)}]^T$ for Rx-1, $\mathbf{s}_2 = [\mathbf{s}_2^{(1)}\ \mathbf{s}_2^{(2)}]^T, \bar{\mathbf{s}}_2 = [\bar{\mathbf{s}}_2^{(1)}\ \bar{\mathbf{s}}_2^{(2)}]^T$ for Rx-2 and $\mathbf{s}_3 = [\mathbf{s}_3^{(1)}\ \mathbf{s}_3^{(2)}]^T, \bar{\mathbf{s}}_3 = [\bar{\mathbf{s}}_3^{(1)}\ \bar{\mathbf{s}}_3^{(2)}]^T$ for Rx-3, where all the symbol vectors have dimension $2N \times 1$ and contain two sub-vectors with the same dimension.

Phase 1

In the first phase, the BS sends signal vectors to two of the three receivers at each time slot, i.e., the symbol vectors sent in the first time slot can be written as,

$$\mathbf{x}(1) = \left[\mathbf{Q}_2(1)\ \ \mathbf{Q}_2^{\perp}(1)\right]\mathbf{s}_1 + \left[\mathbf{Q}_1(1)\ \ \mathbf{Q}_1^{\perp}(1)\right]\mathbf{s}_2,$$

where $\mathbf{Q}_i(t)$ and $\mathbf{Q}_i^{\perp}(t)$ are in the range and the orthogonal space of $\hat{\mathbf{H}}_i(t)$, respectively. The power allocation of each of the symbol vectors are $\mathbb{E}[\|\mathbf{s}_1^{(1)}\|^2] \doteq \mathbb{E}[\|\mathbf{s}_2^{(1)}\|^2] \doteq P^{1-\alpha}$ and $\mathbb{E}[\|\mathbf{s}_1^{(2)}\|^2] \doteq \mathbb{E}[\|\mathbf{s}_2^{(2)}\|^2] \doteq P$. Symmetrically, symbol vectors sent in the following two time slots are given by,

$$\mathbf{x}(2) = \left[\mathbf{Q}_3(2)\ \ \mathbf{Q}_3^{\perp}(2)\right]\bar{\mathbf{s}}_1 + \left[\mathbf{Q}_1(2)\ \ \mathbf{Q}_1^{\perp}(2)\right]\mathbf{s}_3,$$
$$\mathbf{x}(3) = \left[\mathbf{Q}_3(3)\ \ \mathbf{Q}_3^{\perp}(3)\right]\bar{\mathbf{s}}_2 + \left[\mathbf{Q}_2(3)\ \ \mathbf{Q}_2^{\perp}(3)\right]\bar{\mathbf{s}}_3.$$

At the receiver side, the received symbols and the corresponding power allocations in the first time

Common Symbol Vector	Symbol Vector Contained	Generating Time Slot
$\mathbf{c}_1$	$\mathbf{s}_2$	Time Slot 1
$\bar{\mathbf{c}}_1$	$\mathbf{s}_3$	Time Slot 2
$\mathbf{c}_2$	$\mathbf{s}_1$	Time Slot 1
$\bar{\mathbf{c}}_2$	$\bar{\mathbf{s}}_3$	Time Slot 3
$\mathbf{c}_3$	$\bar{\mathbf{s}}_1$	Time Slot 2
$\bar{\mathbf{c}}_3$	$\bar{\mathbf{s}}_2$	Time Slot 3

Table 4.1: Private symbol vector contained in each of the common symbol vector.

slot are given by,

$$\mathbf{y}_1(1) = \underbrace{\mathbf{H}_1(1)\mathbf{Q}_2(1)\mathbf{s}_1^{(1)}}_{P} + \underbrace{\mathbf{H}_1(1)\mathbf{Q}_2^{\perp}(1)\mathbf{s}_1^{(2)}}_{P^{1-\alpha}} + \underbrace{\boldsymbol{\eta}_1}_{P^{1-\alpha}} + \underbrace{\mathbf{n}_1(1)}_{P^0},$$
$$\mathbf{y}_2(1) = \underbrace{\mathbf{H}_2(1)\mathbf{Q}_1(1)\mathbf{s}_2^{(1)}}_{P} + \underbrace{\mathbf{H}_2(1)\mathbf{Q}_1^{\perp}(1)\mathbf{s}_2^{(2)}}_{P^{1-\alpha}} + \underbrace{\boldsymbol{\eta}_2}_{P^{1-\alpha}} + \underbrace{\mathbf{n}_2(1)}_{P^0}.$$

where $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ contain the interference at Rx-1 and Rx-2 respectively and Rx-3 regards all its received signal as interference. Power of the interference term at Rx-1 is given by,

$$\boldsymbol{\eta}_1 = \underbrace{\mathbf{H}_1(1)\mathbf{Q}_1(1)\mathbf{s}_2^{(1)}}_{P^{1-\alpha}} + \underbrace{\mathbf{H}_1(1)\mathbf{Q}_1^{\perp}(1)\mathbf{s}_2^{(2)}}_{P^{1-\alpha}},$$

by which $\mathbb{E}[\|\boldsymbol{\eta}_1\|^2] \doteq P^{1-\alpha}$. Similarly, $\boldsymbol{\eta}_2$ has the same power. At the end of each time slot, the BS will obtain the perfect CSIT of this time slot, which enables it to reconstruct the interference terms, i.e., $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$. In order to make room for privates symbol that will be sent in later phases, each of the reconstructed interference will be digitised into $N(1-\alpha)\log P$ bits to make the digitisation error $\tilde{\boldsymbol{\eta}}_i$ submerged in white noise. After the digitisation process, the digitised version of $\boldsymbol{\eta}_i$ is included in the new common symbol vector $\mathbf{c}_i$ that will be sent in the next phase, where $\mathbb{E}[\|\boldsymbol{c}_i\|^2] \doteq P$.

The symbols received in the other two time slots in the first phase have similar form as the symbols received in the first time slot. In general, the BS reconstructs 6 interference terms and each one will be digitised into $N(1-\alpha)\log P$ bits. The private symbols contained in each of the common symbol vectors and the corresponding generating time slot of the common symbol vectors are listed in Table 4.1.

Phase 2

The second phase also consists of 3 time slots and the symbol vectors sent at each time slot is symmetric. Specifically, in the 4-th time slot, the symbol vectors transmitted are given by,

$$\mathbf{x}(4) = \begin{bmatrix} \mathbf{c}_1 \\ \\ \mathbf{c}_2 \end{bmatrix} + \mathbf{Q}_2^{\perp}(4)\mathbf{s}_1^{(3)} + \mathbf{Q}_1^{\perp}(4)\mathbf{s}_2^{(3)},$$

where $\mathbf{s}_1^{(3)} \in \mathbb{C}^{N\times 1}$ and $\mathbf{s}_2^{(3)} \in \mathbb{C}^{N\times 1}$ are new private symbols for Rx-1 and Rx-2 respectively and $\mathbb{E}[\|\mathbf{s}_1^{(3)}\|^2] \doteq \mathbb{E}[\|\mathbf{s}_2^{(3)}\|^2] \doteq P^{\alpha}$. The corresponding received symbols at the receivers at the 4-th time slot are given by,

$$\begin{aligned}
\mathbf{y}_1(4) &= \underbrace{\mathbf{H}_{1,1}(4)\mathbf{c}_2 + \mathbf{H}_{1,1}(4)\mathbf{c}_1}_{P} + \underbrace{\mathbf{H}_1(4)\mathbf{Q}_2^{\perp}(4)\mathbf{s}_1^{(3)}}_{P^{\alpha}} + \underbrace{\mathbf{H}_1(4)\mathbf{Q}_1^{\perp}(4)\mathbf{s}_2^{(3)}}_{P^0} + \underbrace{\mathbf{n}_1(4)}_{P^0}, \\
\mathbf{y}_2(4) &= \underbrace{\mathbf{H}_{2,1}(4)\mathbf{c}_2 + \mathbf{H}_{2,1}(4)\mathbf{c}_1}_{P} + \underbrace{\mathbf{H}_2(4)\mathbf{Q}_2^{\perp}(4)\mathbf{s}_1^{(3)}}_{P^0} + \underbrace{\mathbf{H}_2(4)\mathbf{Q}_1^{\perp}(4)\mathbf{s}_2^{(3)}}_{P^{\alpha}} + \underbrace{\mathbf{n}_2(4)}_{P^0}, \\
\mathbf{y}_3(4) &= \overbrace{\mathbf{H}_{3,1}(4)\mathbf{c}_2}^{\mathbf{d}_2} + \overbrace{\mathbf{H}_{3,1}(4)\mathbf{c}_1}^{\mathbf{d}_1} + \mathbf{H}_3(4)\mathbf{Q}_{21}^{\perp}(4)\mathbf{s}_1^{(3)} + \mathbf{H}_3(4)\mathbf{Q}_1^{\perp}(4)\mathbf{s}_2^{(3)} + \mathbf{n}_3(4)
\end{aligned}$$

The achievable rate at each receiver to decode combination of the common symbol vectors, i.e., $\mathbf{H}_{i,1}(4)\mathbf{c}_2 + \mathbf{H}_{i,1}(4)\mathbf{c}_1$ is $(1-\alpha)\log P$ bits/s by regarding the all the other terms as noise. After decoding the combination of the common symbol vectors, receivers will be able to remove it from the received signals and decode their private symbols with the remaining interference power being at noise level. After obtaining perfect CSIT at the end of each time slot, the BS will generate new symbols by reconstructing interference terms observed at the receiver which will not use the interference terms. For example, at the 4-th time slot, the BS will generate $\mathbf{d}_2$ and $\mathbf{d}_1$ as $\mathbf{H}_{3,1}(4)\mathbf{c}_2$ and $\mathbf{H}_{3,1}(4)\mathbf{c}_1$, respectively. Note that the terms containing $\mathbf{c}_i$ and $\bar{\mathbf{c}}_i$ are denoted as $\mathbf{d}_i$ and $\bar{\mathbf{d}}_i$, respectively.

The BS will send $\bar{\mathbf{c}}_1$ and $\mathbf{c}_3$ together with new private symbols for Rx-1 and Rx-3. Similarly, the BS will send $\bar{\mathbf{c}}_2$ and $\bar{\mathbf{c}}_3$ together with new private symbols for Rx-2 and Rx-3. At the end of the second phase, the BS has reconstructed $\bar{\mathbf{d}}_1$, $\bar{\mathbf{d}}_2$, $\bar{\mathbf{d}}_3$ and $\mathbf{d}_3$. At the end of the second phase, Rx-1 has decoded $\bar{\mathbf{d}}_2 + \bar{\mathbf{d}}_3$, Rx-2 has decoded $\bar{\mathbf{d}}_1 + \mathbf{d}_3$ and Rx-3 has decoded $\mathbf{d}_1 + \mathbf{d}_2$.

Phase 3

The symbol vectors transmitted the remaining two time slots are given by,

$$\mathbf{x}(7) = u_7 \begin{bmatrix} \bar{\mathbf{d}}_2 + \bar{\mathbf{d}}_3 \\ \mathbf{0} \end{bmatrix} + v_7 \begin{bmatrix} \bar{\mathbf{d}}_1 + \mathbf{d}_3 \\ \mathbf{0} \end{bmatrix} + w_7 \begin{bmatrix} \mathbf{d}_1 + \mathbf{d}_2 \\ \mathbf{0} \end{bmatrix} + \mathbf{c}_7,$$

$$\mathbf{x}(8) = u_8 \begin{bmatrix} \bar{\mathbf{d}}_2 + \bar{\mathbf{d}}_3 \\ \mathbf{0} \end{bmatrix} + v_8 \begin{bmatrix} \bar{\mathbf{d}}_1 + \mathbf{d}_3 \\ \mathbf{0} \end{bmatrix} + w_8 \begin{bmatrix} \mathbf{d}_1 + \mathbf{d}_2 \\ \mathbf{0} \end{bmatrix} + \mathbf{c}_8,$$

where $u_i, v_i, w_i, i = 7, 8$ are shared with all the receivers and the common symbol vector $\mathbf{c}_t, t = 7, 8$, where $\mathbb{E}[\|\mathbf{c}_t\|^2] \doteq P^{\alpha}$, contains two private symbol vectors just as those in the second phase. To keep the scheme symmetric to all the receivers, private symbol vectors contained in $\mathbf{c}_i$ are evenly desired by the three receivers. At the third phase, the received symbols at Rx-1 is given by,

$$\mathbf{y}_i(t) = \mathbf{H}_{i,1}(t) \left[u_t(\bar{\mathbf{d}}_2 + \bar{\mathbf{d}}_3) + v_t(\bar{\mathbf{d}}_1 + \mathbf{d}_3) + w_t(\mathbf{d}_1 + \mathbf{d}_2) + \mathbf{c}_t \right], i = 1, 2, 3,\ t = 7, 8, \quad (4.9)$$

where $\mathbf{H}_{i,1}(t) \in \mathbb{C}^{N \times N}$ is the left half of $\mathbf{H}_i(t)$. By regarding the common symbol vector $\mathbf{c}_t$ as noise, receivers will be able to decode sum of all the other terms at the rate of $(1-\alpha)\log P$ bits/s. Note that sum of all the other terms contains $(1-\alpha)\log P$ bits. After that, the private symbol can be decoded at the rate of $N\alpha \log P$ bits/s.

Remark: Scheme $\mathcal{X}_9$ is hard to extend to IC. The major reason behind this is that in the third phase, the BS is supposed to transmit special combination of order-3 symbols. This special combination of order-3 common symbol can only be sent by a transmitter which knows all the private symbols. For example, in (4.9), $\bar{\mathbf{d}}_2$ and $\bar{\mathbf{d}}_3$ have the same coefficient u_t so that Rx-1 can remove both of them from the received signal in the third phase. This is because Rx-1 only have knowledge of $\bar{\mathbf{d}}_2 + \bar{\mathbf{d}}_3$ but not $\bar{\mathbf{d}}_2$ and $\bar{\mathbf{d}}_3$ independently. If considering IC, only Tx-2 can generate $\bar{\mathbf{d}}_3$ and only Tx-3 can generate $\bar{\mathbf{d}}_2$. Therefore, receivers will always receive $\bar{\mathbf{d}}_2$ and $\bar{\mathbf{d}}_3$ with the different coefficients because of the channel fading.

By regarding all the new private symbols in Phase 2 and Phase 3 as noise, the backwards decoding method is exactly the same as that stated in [8]. Therefore, all the symbols contained in $\mathbf{s}_1$, $\bar{\mathbf{s}}_1$, $\mathbf{s}_2$, $\bar{\mathbf{s}}_2$, $\mathbf{s}_3$, $\bar{\mathbf{s}}_3$ can be successfully decoded. The new private symbol vectors can be decoded using SIC

method.

Degrees of Freedom Calculation

During the last two phases, the new private symbol vectors are transmitted at the power level P^{α}, which means $N\alpha \log P$ bits of information can be successfully decoded from each of them. During the three time slots of the second phase, two symbol vectors are sent to each of the receivers, thus each receiver obtain $2N\alpha$ DoF. At the two time slots of the last phase, the BS evenly allocates the 4 symbols to the three receivers, which means each receiver will obtain $N\frac{4\alpha}{3}$ DoF. The DoF achieved by decoding the private symbols in the last two phases is given by,

$$2N\alpha + N\frac{4\alpha}{3} = N\frac{10\alpha}{3}.$$

For each of the receiver, there are another two symbols successfully decoded, i.e., the private symbol vectors sent in the first phase. As the two symbol vectors have power P and $P^{1-\alpha}$, respectively. The total rate achieved at each receiver is $N(1-\alpha+1)\log P = N(2-\alpha)\log P$.

Adding the DoF achieved when decoding all the private symbols in the 8 time slots together, the per time slot-DoF achieved by Rx-1 are,

$$\frac{2N(2-\alpha) + N\frac{10\alpha}{3}}{8} = N\frac{3+\alpha}{6}.$$

Symmetrically, Rx-2 and Rx-3 obtain the same DoF.

4.7 Simulation Results

The numerical results for Scheme $\mathcal{X}_1$ and Scheme $\mathcal{X}_3$ are shown in this section. For Scheme $\mathcal{X}_3$ which consists of an infinite number of phases, it is approximated with 20 phases. All of the points in the following two figures are averaged over 500 channel realisations.

Figure 4.5 illustrates the sum rate comparison between the Scheme $\mathcal{X}_1$, the TDMA scheme and the ZF scheme. The line with $\alpha = 0, \beta = 0$ represents the rates without any CSIT and the line with $\alpha = 1, \beta = 1$ represents the rates with perfect current and delayed CSIT. From this figure, it can be seen that the TDMA line overlaps with the line when $\alpha = 0, \beta = 0.5$ because both achieve

one DoF in total. Importantly, as discussed in Section 4.4, Scheme $\mathcal{X}_1$ loses to the TDMA scheme when the quality of current and delayed CSIT are poor, which can be summarised from the fact that the TDMA line outperforms the line with $\alpha = 0, \beta = 0$. Note that the ZF line is drawn with a variance of channel estimation error $P^{-\alpha}, \alpha = 0.3$ and therefore, the ZF line achieves 0.6 DoF in total. The ZF line and the line with $\alpha = 0, \beta = 0$, which achieves 0.66 DoF in total, tend to have a similar slope at the high SNR range, which suggests that approximately the same DoF are achieved.

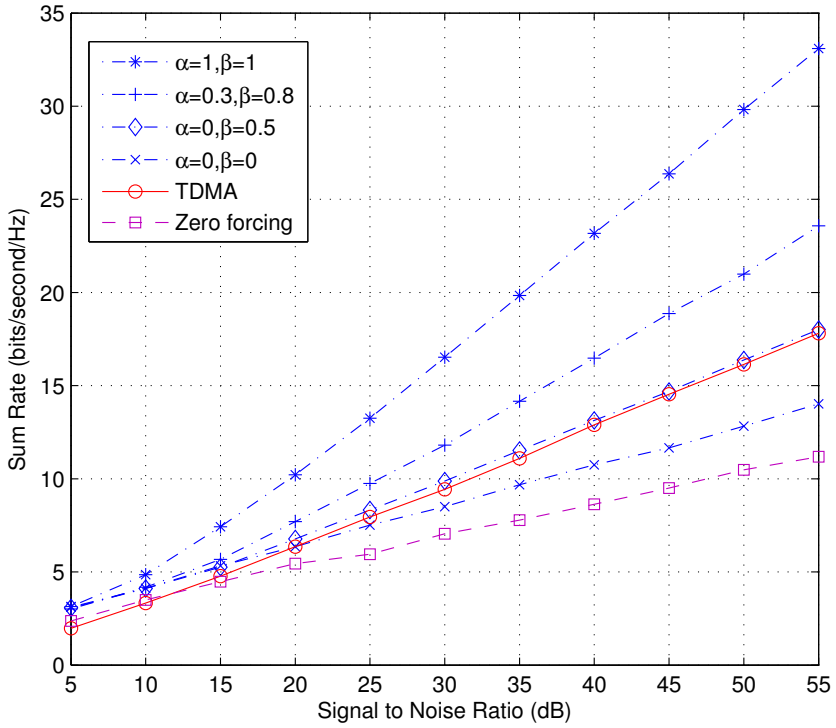


Figure 4.5: Sum rate achieved by Scheme $\mathcal{X}_1$ compared with the TDMA scheme and the ZF scheme.

The simulation results of Scheme $\mathcal{X}_3$ are illustrated in Figure 4.6 for $M = 2, N = 1$ in order to make a fair comparison with Scheme $\mathcal{X}_1$. The number of phases considered in this simulation is 20. The upper most line represents the situation with perfect current and delayed CSIT. A loss is observed when $\alpha = 0.6, \beta = 1$ (perfect delayed CSIT and imperfect current CSIT), which is caused by a decrease of α. The line $\alpha = 0.6, \beta = 0.8$ (imperfect current and delayed CSIT) overlaps with the line $\alpha = 0.6, \beta = 1$ because in both situation, $\beta \geq \frac{1+2\alpha}{3}$ and, therefore, changing β does not affect the DoF region as predicted in Section 4.4. Different from Figure 4.5, the line $\alpha = 0, \beta = 0$ overlaps with the TDMA line because both achieve *one* DoF as shown in Figure 4.2. In other words, Scheme $\mathcal{X}_3$ performs equivalently to the TDMA scheme in the worst

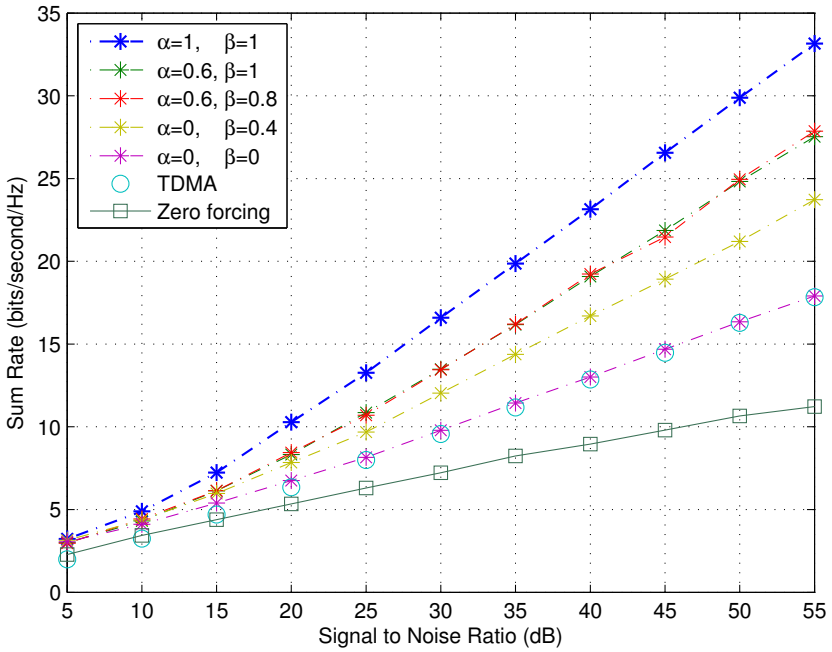


Figure 4.6: Sum rate achieved by Scheme $\mathcal{X}_3$ compared with the TDMA scheme and the ZF scheme.

case. The line with $\alpha = 0, \beta = 0.4$ outperforms the TDMA scheme, which is due to the benefit of having knowledge of delayed CSIT. In this situation, $\beta < \frac{1+2\alpha}{3}$, so the DoF pair achieved is $(\frac{1+\beta}{2}, \frac{1+\beta}{2})$ as stated in Theorem 4. Finally, the lowest line corresponding to the ZF scheme ($\alpha = 0.3$) achieves 0.6 DoF in total and, therefore, underperforms all of the other lines, which at least achieve a single DoF.

4.8 Summary

This chapter proposed the DoF regions and the corresponding achievability schemes in the time-correlated channels using the current and delayed CSIT. In total, nine schemes were provided to achieve different vertices in the following four DoF regions: 1) 2-user and 2-antenna transmitter MISO IC with imperfect current and delayed CSIT; 2) 2-user and $2N$-antenna transmitter MIMO (N receiver antennas) IC with imperfect current and delayed CSIT; 3) K-user and KN-antenna BS MIMO (N receiver antennas) BC with imperfect current CSIT and perfect delayed CSIT; 4) 3-user and $2N$-antenna BS MIMO (N receiver antennas) BC with imperfect current CSIT and perfect delayed CSIT. All of the DoF regions considering MIMO in this chapter were N times linear expansions of the corresponding DoF regions when $N = 1$. Hence, the following

conclusions are given for the case when $N = 1$:

- **2-user and 2-antenna IC with imperfect current and delayed CSIT:** A three-phase scheme, Scheme $\mathcal{X}_1$, was proposed first and achieved $\frac{1+\alpha+\beta}{3}$ DoF at each receiver in $\mathcal{S}^{(1)}_{MAT}$ with imperfect current and delayed CSIT. It could be seen that when $\alpha+\beta < \frac{1}{2}$, Scheme $\mathcal{X}_1$ underperformed the TDMA scheme, which required no CSIT. In other words, the CSIT was harmful to the DoF region if it had low quality, which was contrary to common knowledge. In order to tackle this issue, an advanced scheme was illustrated, Scheme $\mathcal{X}_3$, which was an extension of the work in [66]. When $\alpha = \beta = 0$, each receiver achieved $\frac{1}{2}$ DoF and when $\alpha = \beta = 1$, each receiver achieved *one* DoF. That is to say, the CSIT was always useful. It was also shown that Scheme $\mathcal{X}_3$ strictly outperformed Scheme $\mathcal{X}_1$ once $\beta < 1$. For Scheme $\mathcal{X}_3$, there was no need to increase the quality of the delayed CSIT once $\beta = \frac{1+2\alpha}{3}$.

- **K-user and K-antenna BS with imperfect current CSIT:** How to achieve the DoF outer bound proposed in [67] with K-user and K-antenna BS using imperfect current CSIT and perfect delayed CSIT was an open problem. Firstly, Scheme $\mathcal{X}_5$ was proposed for the case of $K = 3$ and the corresponding 3-dimensional DoF region was plotted. This scheme was extended to the case with general number of users, where Scheme $\mathcal{X}_7$ was proposed achieving $\frac{K(1-\alpha)}{1+\frac{1}{2}+\frac{1}{3}+\cdots\frac{1}{K}} + K\alpha$ DoF in total, which was exactly the DoF outer bound in [67]. With varying α, Scheme $\mathcal{X}_7$ compromised between the MAT scheme and the ZF beamforming scheme with perfect current CSIT. Moreover, Scheme $\mathcal{X}_8$ was proposed which gave one receiver *one* DoF and all of the other receivers α DoF. In total, Scheme $\mathcal{X}_8$ achieved K DoF when $\alpha = 1$. Because special common symbol vectors were included, which required that the transmitter had knowledge of all of the private symbols, Scheme $\mathcal{X}_5$ and Scheme $\mathcal{X}_7$ could not be extended to IC.

- **3-user and 2-antenna BS with imperfect current CSIT:** Specifically, a 3-user BC where $M = 2$ was considered in the presence of imperfect current CSIT and perfect delayed CSIT. Firstly, equations of all the surfaces of the 3-dimensional DoF polygon was illustrated and the polygon was proven to be convex. Afterwards, Scheme $\mathcal{X}_9$ was proposed which achieved $\frac{3+\alpha}{6}$ DoF at each receiver and the scheme was consistent with two well known results, i.e., when $\alpha = 0$, in total, $\frac{3}{2}$ DoF were achieved as obtained by the MAT scheme and when $\alpha = 1$, 2 DoF were achieved as obtained by the ZF beamforming scheme with perfect current CSIT. That is to say, the current CSIT is always beneficial. Similarly, because special common symbol vectors were required, Scheme $\mathcal{X}_9$ could not be extended to IC.

Chapter 5

Interference Cancellation in Poisson Distributed Wireless Networks

5.1 Introduction

Large-scale wireless networks such as cellular networks and MANETs have attracted a considerable amount of attention from researchers. Given the fast development of such wireless networks, it is natural to investigate their capacities in the presence of random distribution. The authors in [68–70] studied the capacity of MANETs with consideration of the path-loss effect where transmitters formed a deterministic pattern. The authors in [10] studied the wireless networks with ALOHA-type transmission (a random access protocol) and, more importantly, a brand-new metric for the area spectral efficiency, called *transmission capacity*, was proposed. Note that transmission capacity was introduced in Section 2.3.11. With the development of stochastic geometry technique, a number of work studied transmission capacity in wireless networks such as [23, 71–73].

Most of the previous work studying transmission capacity in Poisson wireless networks commonly assumes a single-antenna setting and without any interference cancellation technique. An exception is [23], in which transmission capacity in a single-antenna PPP network was studied using SIC in the absence of fading effect. As MIMO-based techniques are of great importance in all kinds of wireless network with the capability to combat fading and reduce interference [74, 75], how to evaluate its performance in distributed wireless networks is of great interest. Hence, in Section 5.3, the average SINR as well as transmission capacity achieved by the ZF receiver was investigated in a PPP wireless network. Unlike [23], Rayleigh fading, multiple-antenna setting and diversity combining are considered, which make the distributions of the power of received signals more complex. After the work in Section 5.3 was accomplished, the authors in [9] also investigated the transmission capacity achieved by the partially ZF receiver in a multiple-antenna PPP wireless network, in which the effectiveness of CSIT and the trade-off between cancelling

interference and increasing the multiplexing gain were studied.

Stochastic geometry-based approaches successfully characterised the randomness of node distribution in 2-dimensional planes. However, most future wireless networks will be packet-switched with bursty data traffic, requiring frequent changes in the number of active users [76]. Hence, one of the objectives of the remaining two sections is to model both the random node distribution and bursty data traffic in wireless networks. Bursty data traffic has also been considered in [77], in which the authors studied the achievable sum rates in an IA network by the active transmitter-receiver pairs as well as secondary ones. Note that the *secondary node* is the one that wants to join the network. Unlike [77], where all the secondary transmitter-receiver pairs joined the network simultaneously and the path-loss effect was not considered, this chapter assumes all of the nodes join the network independently in the presence of the path-loss effect. It is worth mentioning that, in general, a high overhead cost is associated with finding the IA beamforming solution. Therefore, once the active transmitters have successfully aligned their interference, it is desirable to retain their alignment status.

In this chapter, multiple-antenna PPP wireless networks are considered. Specifically, Section 5.4 studies the average SINR and the transmission capacity achieved by the ZF receiver which can design the *CSI area* and feasibly eliminate all inside interference. Note that in the CSI area of a receiver, it has instantaneous and perfect current CSI of the links it is connected to. Next, Section 5.4, which is a bridge between Section 5.3 and Section 5.5, considers that the ZF receiver cannot always feasibly eliminate all the interference in a fixed CSI area when joining the network. In other words, it studies the influence of feasibility condition to the transmission capacity achieved by the secondary receiver in the presence of finite CSI area. In the condition that the CSI area is infinitely large, this section reduces to the work in [9]. Finally, Section 5.5 studies the performance of the ZF receiver with random node distribution and random data traffic, wherein the random data traffic is modelled using queueing theory as in [78–80] and the random node distribution is modelled using stochastic geometry. In this section, IA beamforming is considered and the evolving success probability at the receiver is derived using the queueing model. The feasibility condition of the ZF receiver is also incorporated with the queueing model to determine the probability that it is allowed to join the network. Unlike [9], in which interference from a few nearest transmitters is eliminated and bounds of transmission capacity were derived, this chapter successfully obtains the accurate expression (not bounds) of the transmission capacity where the interference-free region of each receiver is a ball with a fixed radius.

The remainder of this chapter is organised as follows. Section 5.2 builds the system model for all the following sections. Section 5.3 studies the average SINR and transmission capacity achieved by the ZF receiver that can design the CSI area and eliminate all the inside interference. Section 5.4 studies the influence of the feasibility condition of the ZF receiver to the achievable transmission capacity. Section 5.5 evaluates the performance achieved by the ZF receiver in a Poisson wireless network where transmitter-receiver pairs join and leave independently. The simulation results and summaries are in Section 5.6 and Section 5.7, respectively.

5.2 System Model

Consider a multiple-antenna Poisson wireless network where the transmitter set $\Phi_T = \{T_i | i \in \mathbb{N}\}$ forms PPP with intensity λ and the i-th receiver is denoted as R_i. Note that T_i or R_i are utilised to represent the location of the corresponding transmitter or receiver. Each receiver is assumed to have perfect CSI to all of the transmitters in its CSI area and the CSI areas are balls with radius r_c. By assuming that the inactive receivers are listening, all the secondary receivers have access to the CSI of the active and secondary transmitters in their CSI areas. The receivers exploit ZF approach to eliminate the interference from the transmitters inside its CSI area. The effective range in which interference is cancelled is called the ZF area and its radius is denoted as r_z. Straightforwardly, $r_z \leq r_c$. For each receiver, interference outside the ZF area is treated as shot noise as stated in [37].

The signal model in detail for the typical transmitter-receiver pair (T_0, R_0) is describe in detail. The performance achieved by R_0 also represents the average performance of all the receivers. The CSI area of R_0 is denoted as $\mathcal{A}_c$ and the ZF area is denoted as $\mathcal{A}_z$. The post-processing signal received at R_0 is given by,

$$\hat{\mathbf{y}}_0 = l_{00}^{-\varsigma/2}\mathbf{U}_0^H\mathbf{H}_{00}\mathbf{V}_0\mathbf{s}_0 + \mathbf{U}_0^H \sum_{\substack{T_i \in \Phi_T/\mathcal{A}_z, \\ T_i \neq T_0}} l_{0i}^{-\varsigma/2}\mathbf{H}_{0i}\mathbf{V}_i\mathbf{s}_i + \mathbf{n}_0,$$

where

- $\mathbf{s}_i \in \mathbb{C}^{\bar{d}\times 1}$ is the symbol vector sent by T_i;
- $\mathbf{H}_{ij} \in \mathbb{C}^{N\times M}$ is the channel matrix from T_j to R_i with i.i.d. $\mathcal{CN}(0,1)$ entries;
- $\mathbf{V}_i \in \mathbb{C}^{M\times \bar{d}}$ is the beamforming matrix at T_i and its columns have unit norm;

- $\mathbf{U}_0 \in \mathbb{C}^{N\times\bar{d}}$ is the ZF receiving matrix at Rx-0;
- $\mathbf{n}_i \in \mathbb{C}^{\bar{d}\times 1}$ is the zero-mean unit-variance additive white noise observed by R_i,

where $\bar{d}$ is the number of data streams, each receiver has N antennas and each transmitter has M antennas. Because interference is dominant in this chapter, the additive white noise is ignored by default.

At R_0, in order to decode the desired symbol in the q-th data stream, the decorrelator $\mathbf{f}_q$, which is defined in Section 2.2.3, should be used so that the desired signal power is maximised. As a result, the desired signal and the interference signal at the q-th data stream can be expressed as $l_{00}^{-\varsigma/2}\mathbf{f}_q^H\mathbf{H}_{00}\mathbf{V}_0\mathbf{s}_0$ and $\mathbf{f}_q^H\sum_{T_i\in\Phi_T/\mathcal{A}_z} l_{0i}^{-\varsigma/2}\mathbf{H}_{0i}\mathbf{V}_i\mathbf{s}_i$, respectively. The signal power and overall interference power at the q-th data stream of R_0 are given by,

$$s_{0,q} = l_{00}^{-\varsigma}\|\mathbf{f}_q^H\mathbf{H}_{00}\mathbf{V}_0\|^2, \quad I_q = \|\mathbf{f}_q^H\sum_{T_i\in\Phi_T/\mathcal{A}_z} l_{0i}^{-\varsigma/2}\mathbf{H}_{0i}\mathbf{V}_i\mathbf{s}_i\|^2,$$

In the case of Rayleigh fading and R_0 performing single stream decoding as in [9], R_0 can exploit $N-\bar{d}_I-\bar{d}+1$ SRDoF to decode each desired symbol and thus $s_{0,q}$ becomes chi-squared distributed with $2(N-\bar{d}_I-\bar{d}+1)$ DoF, where $\bar{d}_I$ is the dimension of the interference subspace at R_0. The corresponding complementary CDF of the desired signal power is $\tilde{F}_\Gamma(x) = \frac{\Gamma(N-\bar{d}_I-\bar{d}+1,x)}{\Gamma(N-\bar{d}_I-\bar{d}+1)}$. If the pre-specified SIR is chosen as θ, the outage probability at the q-th data stream of R_0 can be written as the following equation:

$$\epsilon = 1 - \mathbb{E}_{I_q}\left[\frac{\Gamma(N-\bar{d}_I-\bar{d}+1,\theta l_{00}^{\varsigma}I_q)}{\Gamma(N-\bar{d}_I-\bar{d}+1)}\right]. \tag{5.1}$$

A more specific system model will be given for each section.

5.3 ZF Receiver in A Static Network

In this section, $\bar{d} = 1$ is assumed and R_0 can design the radius of the CSI area and feasibly eliminate all interference from the transmitters in the CSI area, i.e., $r_z = r_c$. The transmitter T_0, on the other hand, does not perform any beamforming to increase the desired signal power at R_0. The average SINR as well as the upper and lower bounds of transmission capacity in the Rayleigh fading model are derived.

5.3.1 Nearest Neighbour Distance

The CDF of the distance from R_0 to the nearest transmitter in PPP is derived in [81] and it can be extended to the CDF of the distance to the k-th nearest transmitter, which is denoted as $F_k(l)$. Intuitively, the CDF can be written as,

$$F_k(l) = 1 - \lim_{r_\xi \to 0} \left\{ \sum_{i=0}^{k-1} \Pr(\Phi(b(R_0, l) \backslash b(R_0, r_\xi)) = i) \right\}.$$

Taking the Poisson distribution into consideration, the following equations can be obtained:

$$F_k(l) = 1 - \lim_{r_\xi \to 0} \left\{ \sum_{i=1}^{k-1} \frac{(\lambda\pi(l^2 - r_\xi^2))^i e^{-\lambda\pi(l^2 - r_\xi^2)}}{i!} \right\} = 1 - \sum_{i=1}^{k-1} \frac{(\lambda\pi l^2)^i e^{-\lambda\pi l^2}}{i!}.$$

The expected value of the distance between R_0 and the k-th nearest transmitter, denoted as l_k, can be expressed as,

$$\mathbb{E}[l_k] = \int_0^\infty l \mathrm{d}F_k(l) = \sum_{i=1}^{k-1} \frac{2(\lambda\pi)^{i-1}}{(i-1)!} \int_0^\infty \left(\lambda\pi l^{2i-1} - (i-1) l^{2i-3}\right) e^{-\lambda\pi l^2} l \mathrm{d}l.$$

Due to the following equation:

$$\int_0^\infty l^{2i} e^{-\lambda\pi l^2} \mathrm{d}l = \frac{(2i-1)!!}{2(2\lambda\pi)^i \sqrt{\lambda}},$$

the expected distance from R_0 to the k-th nearest transmitter can be simplified as,

$$\mathbb{E}[l_k] = \frac{(2k-1)!!}{2^k (k-1)! \sqrt{\lambda}}, \tag{5.2}$$

where $(\cdot)!!$ is the double factorial. In the remainder of this section, $r_z = \mathbb{E}[l_{\hat{k}}]$ is assumed, where $\hat{k}$ is a constant number called the *coverage factor*. The coverage factor determines the radius of the CSI area so that $\hat{k}$ number of transmitters is expected to be included. In the remainder of this section, it is assumed that the actual number of transmitters in the CSI area is $\hat{k}$.

5.3.2 Average Signal-to-Interference-Noise Ratio

If a coverage factor $\hat{k}$ is selected, then R_0 will spend $\hat{k}$ SRDoF to cancel the interference from transmitters in the ZF area. Using diversity combining, the effective SRDoF to decode its own

private symbol is $N - \hat{k}$. Hence, the expected desired signal power can be expressed as,

$$\mathbb{E}[s_{0,q}] = l_{00}^{-\varsigma}(N - \hat{k}).$$

Because all the interference inside the ZF area has been eliminated, the expected value of the remaining interference power can be obtained by the following equation:

$$\mathbb{E}[I_q] = \mathbb{E}[h] \int_{r_z}^{\infty} 2\pi\lambda x x^{-\alpha} \mathrm{d}x = \frac{2\pi\lambda r_z^{2-\varsigma}}{\varsigma - 2},$$

where h is the fading coefficient with expected value *one*. Hence, the average SINR can be expressed as,

$$\mathbb{E}\left[\frac{s_{0,q}}{I_q + n_0}\right] = \frac{l_{00}^{-\varsigma}(N - \hat{k})}{\frac{2\pi\lambda r_z^{2-\varsigma}}{\varsigma - 2} + n_0}.$$

where n_0 is the white noise power at R_0 and all of the symbols are sent with unit power. Taking (5.2) into consideration, the average SINR can be expressed as

$$\mathbb{E}\left[\frac{s_{0,q}}{I_q + n_0}\right] \approx \frac{l_{00}^{-\varsigma}(N - \hat{k})(\varsigma - 2)}{2\pi\lambda \left(\frac{(2\hat{k}-1)!!}{2^{\hat{k}}(\hat{k}-1)!\sqrt{\lambda}}\right)^{2-\varsigma} + n_0(\varsigma - 2)}, \tag{5.3}$$

The approximation comes from the fact that the actual number of point inside a ball with radius $\mathbb{E}[l_k]$ may differ from k and the average SINR is not a linear function of $\hat{k}$. This expression actually delivers the average SINR achieved at R_0 when $\hat{k}$ nearest interferers are cancelled.

5.3.3 Transmission Capacity with Rayleigh Fading

If $\hat{k} = N - 1$ is selected, i.e., R_0 can only exploit *one* SRDoF to decode its desired symbol, in the case of Rayleigh fading, the desired signal power becomes exponentially distributed. The outage probability is given by (5.1) and the distance between T_0 and R_0 is *one*. Therefore, the success probability, which equals $1 - \epsilon$, can be transferred using the following equations,

$$\begin{aligned} 1 - \epsilon &= \mathbb{E}\left[e^{-sI_q}\right]\big|_{s=\theta} \\ &= \exp\left(\mathbb{E}_h\left[2\pi\lambda \int_{r_z}^{\infty} x(e^{-\theta h x^{-\varsigma}} - 1)dx\right]\right). \\ &= \exp(\lambda \mathbb{E}_h[\mathcal{Q}_5]), \end{aligned} \tag{5.4}$$

where h is a exponentially distributed random variable with unit mean and,

$$\mathcal{Q}_5 = \pi r_z^2(1-\exp(-h\theta r_z^{-\varsigma})) - \pi(h\theta)^{2/\varsigma}\gamma(h\theta r_z^{-\varsigma}, 1-2/\varsigma),$$

where $\gamma(\cdot,\cdot)$ is the lower incomplete gamma function. In this way, the maximum transmitter intensity to guarantee outage probability ϵ is given by,

$$\lambda^{\epsilon} = \frac{\ln(1-\epsilon)}{\mathbb{E}_h[\mathcal{Q}_5]}.$$

Take this value as λ^{ϵ} in (2.6), the transmission capacity can be rewritten as,

$$\mathcal{C} = \frac{R(1-\epsilon)\ln(1-\epsilon)}{\mathbb{E}_h[\mathcal{Q}_5]}. \tag{5.5}$$

Lower Bound of Transmission Capacity

It is straightforward to prove that $\mathcal{Q}_5$ is a concave function in terms of h. Hence, based on the Jensen's inequality,

$$\begin{aligned}\mathbb{E}_h[\mathcal{Q}_5] \geq& \pi r_z^2(1-\exp(-\theta r_z^{-\varsigma})) - \pi\theta^{2/\varsigma}\gamma(\theta r_z^{-\varsigma}, 1-2/\varsigma)\\ \geq& \pi r_z^2\big((1-\exp(-\theta r_z^{-\varsigma})) - \theta^{2/\varsigma}\Gamma(1-2/\varsigma)\big) \qquad (5.6a)\end{aligned}$$

Inequality (5.6a) is based on the fact that $\Gamma(t) \geq \gamma(x,t)$. Taking (5.6a) into (5.5), a lower bound of transmission capacity is obtained as,

$$\mathcal{C}_{lb} = \frac{R(1-\epsilon)\ln(1-\epsilon)}{\pi r_z^2\big((1-\exp(-\theta r_z^{-\varsigma})) - \theta^{2/\varsigma}\Gamma(1-2/\varsigma)\big)},$$

where $\ln(\cdot)$ denotes the natural logarithm.

Upper Bound of Transmission Capacity

On the other side,

$$\begin{aligned}\mathbb{E}_h[\mathcal{Q}_5] \leq& \mathbb{E}[\pi r_z^2(1-\exp(-h\theta r_z^{-\varsigma})) - \pi(h\theta)^{2/\varsigma}\gamma(h\theta r_z^{-\varsigma}, 1)] \qquad (5.7a)\\ =& \pi\mathbb{E}[r_z^2(1-\exp(-h\theta r_z^{-\varsigma})) - (h\theta)^{2/\varsigma}(1-\exp(-h\theta r_z^{-\varsigma}))] \qquad (5.7b)\\ =& \pi\left(r_z^2 - \Gamma(1+2/\varsigma)\theta^{2/\varsigma} - \frac{r_z^2}{\theta r_z^{-\varsigma}+1} + \frac{\theta^{2/\varsigma}\Gamma(1+2/\varsigma)}{(\theta r_z^{-\varsigma}+1)^{2/\varsigma+1}}\right) \qquad (5.7c)\end{aligned}$$

where the inequality (5.7a) is valid for relatively small r_z, e.g., $r_z \leq 2$, and equation (5.7b) is based on $\gamma(x,1) = 1 - e^{-x}$. Taking (5.7c) into (5.5), an upper bound of transmission capacity can be obtained as the following demonstrates:

$$\mathcal{C}_{ub} = \frac{\pi^{-1} R(1-\epsilon)\ln(1-\epsilon)}{r_z^2 - \Gamma(1+2/\varsigma)\theta^{2/\varsigma} - \frac{r_z^2}{\theta r_z^{-\varsigma}+1} + \frac{\theta^{2/\varsigma}\Gamma(1+2/\varsigma)}{(\theta r_z^{-\varsigma}+1)^{2/\varsigma+1}}}.$$

5.4 Influence of the Feasibility Condition of the ZF Receiver

In this section, it is assumed that the transmitter-receiver pair (T_0, R_0) is joining the network and R_0 actively provides all of the variables to eliminate the interference as stated in Section 5.1.

Function $f_n(N, \bar{d})$ is defined as the number of transmitters the interference from which can be eliminated by R_0. If there are fewer than $f_n(N, \bar{d})$ transmitters in the CSI area of R_0, it can eliminate all of the inside interference. This case is called the *perfect ZF condition*. In the perfect ZF condition, $r_z = r_c$ and the following equation will be satisfied:

$$\mathbf{U}_0^H \mathbf{H}_{0i} \mathbf{V}_i = 0, \;\; \forall T_i \in b(R_0, r_z).$$

When there are more than $f_n(N, \bar{d})$ transmitters in $b(R_0, r_z)$, the secondary receiver can not feasibly cancel all the interferences in $b(R_0, r_z)$. This situation is referred to as the *imperfect ZF condition* and the following equation should be satisfied:

$$\mathbf{U}_0^H \mathbf{H}_{0i} \mathbf{V}_i = 0, \;\; i = 1, 2, 3, \ldots f_n(N, \bar{d}).$$

where indices are allocated to transmitters according to their distances to R_0. In the imperfect ZF condition, the radius of ZF area $\mathcal{A}_z$ becomes a function of N, $\bar{d}$ and λ.

5.4.1 Feasibility Condition

As stated in [35], receiver that is equipped with N antennas and receiving $\bar{d}$ data streams can provide $\bar{d}(N - \bar{d})$ variables to eliminate interference and in order to eliminate interference from one transmitter, $\bar{d}^2$ equations are required. As stated before, all of the variables to eliminate the interference is provided by R_0, i.e., the total number of variables is $\bar{d}(N - \bar{d})$. To eliminate

interference from $f_n(N,\bar{d})$ transmitters, there are $\bar{d}^2 f_n(N,\bar{d})$ equations in total. According to the Bezout's theorem, the total number of variables should be no smaller than the total number of equations in order to make the system proper. Therefore, $f_n(N,\bar{d})$ can be expressed as,

$$f_n(N,\bar{d}) \times \bar{d}^2 \leq \bar{d}(N-\bar{d}) \Longleftrightarrow f_n(N,\bar{d}) = \left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor.$$

For a secondary receiver, the perfect ZF condition to join the network can be interpreted as the condition that there are no more than $\left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor$ transmitters in its CSI area. The probability of perfect ZF condition for a secondary receiver can be expressed as,

$$\Pr(\text{pZF}) = \Pr\left(\Phi_T(|b(R_0,r_z)|) \leq f_n(N,\bar{d})\right) = \sum_{k=0}^{f_n(N,\bar{d})} \frac{e^{-\lambda\pi r_z^2}(\lambda\pi r_z^2)^k}{k!} = \frac{\Gamma\left(\left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor + 1, \lambda\pi r_z^2\right)}{\Gamma\left(\left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor + 1\right)},$$

where $\Gamma(\cdot,\cdot)$ is the upper incomplete gamma function. With the probability of the feasibility condition, the overall transmission capacity achieved by secondary receiver can be averaged over the two conditions: $\Pr(\text{pZF})\mathcal{C}_p + (1-\Pr(\text{pZF}))\mathcal{C}_i$, where $\mathcal{C}_p$ and $\mathcal{C}_i$ are the transmission capacities assuming all of the secondary receivers are in perfect ZF condition and imperfect ZF condition, respectively.

5.4.2 Perfect ZF Condition

In perfect ZF condition, interference within $b(R_0,r_z)$ is completely eliminated, then effective interference is only from area outside r_z. Note that $r_z = r_c$ in this case. Hence, the corresponding outage probability can be expressed as,

$$\epsilon_p = \Pr\left(\frac{s_{0,q}}{I_p} < \theta\right) = \Pr\left(\frac{s_{0,q}}{\sum_{T_i \in \mathbb{R}^2/b(R_0,r_z)} l_{0i}^{-\varsigma} \left\|\mathbf{f}_q^H \mathbf{H}_{0i} \mathbf{V}_i\right\|^2} < \theta\right),$$

where $s_{0,q}$ is considered as a constant in this section.

Lower Bound of Transmission Capacity

Based on Markov's inequality and Campbell's theorem [72],

$$\epsilon_p = \Pr\left(\sum_{T_j \in \mathbb{R}^2/b(R_0,r_z)} l_{0i}^{-\varsigma} \left\|\mathbf{f}_q^H \mathbf{H}_{0i} \mathbf{V}_i\right\|^2 > \frac{s_{0,q}}{\theta}\right)$$

$$\leq \frac{\theta}{s_{0,q}}\mathbb{E}\left[\sum_{T_i\in\mathbb{R}^2/b(R_0,r_z)} l_{0i}^{-\varsigma}\left\|\mathbf{f}_q^H\mathbf{H}_{0i}\mathbf{V}_i\right\|^2\right]$$
$$= \frac{\theta}{s_{0,q}}\int_{r_z}^{\infty} x^{-\varsigma}\mathbb{E}[h]\lambda 2\pi x dx.$$

where h denotes the fading coefficient. In the case of Rayleigh fading channel, h is chi-squared distributed with $2\bar{d}$ DoF. Thus, $\mathbb{E}[h] = \bar{d}$. The upper bound of outage probability with perfect ZF condition is given by,

$$\epsilon_p^{up} = \frac{2\pi\lambda\theta\bar{d}r_z^{2-\varsigma}}{(\varsigma-2)s_{0,q}}. \tag{5.8}$$

The corresponding lower bound of the transmission capacity with perfect ZF condition is given by,

$$\mathcal{C}_p^{lb} = \frac{R(1-\epsilon)(\varsigma-2)\epsilon s_{0,q}r_z^{\varsigma-2}}{2\pi\theta}.$$

Upper Bound of Transmission Capacity

A lower bound of the outage probability can be obtained by considering part of the interference. Specifically, only the interference caused by transmitters within distance r_z to r_m ($r_m > r_z$) of R_0 is considered. This part of interference power is denoted as I_r and satisfies the following equation:

$$\mathbb{E}\left[\frac{1}{I_r}\right] = \frac{1}{(\bar{d}-1)\int_{r_z}^{r_m} 2\pi\lambda x^{-\varsigma}xdx} = \frac{2-\varsigma}{2\pi\lambda(\bar{d}-1)(r_m^{2-\varsigma}-r_z^{2-\varsigma})}.$$

With Markov's inequality,

$$1-\epsilon_p \leq \Pr\left(\frac{1}{I_r} > \frac{\theta}{s_{0,q}}\right) \leq \mathbb{E}\left[\frac{1}{I_r}\right]\frac{s_{0,q}}{\theta} = \frac{(2-\varsigma)s_{0,q}}{2\pi\lambda\theta(\bar{d}-1)(r_m^{2-\varsigma}-r_z^{2-\varsigma})}.$$

Therefore, a lower bound of the outage probability in perfect ZF condition can be expressed as,

$$\epsilon_p^{lb} = 1 - \frac{(2-\varsigma)s_{0,q}}{2\pi\lambda\theta(\bar{d}-1)(r_m^{2-\varsigma}-r_z^{2-\varsigma})}.$$

The corresponding upper bound of the transmission capacity is given by,

$$\mathcal{C}_p^{ub} = \frac{(1-\epsilon)R\bar{d}(2-\varsigma)s_{0,q}}{2\pi\theta(\bar{d}-1)(r_m^{2-\varsigma}-r_z^{2-\varsigma})}.$$

5.4.3 Imperfect ZF Condition

In the imperfect ZF condition, R_0 can not eliminate all the interference inside $\mathcal{A}_c$. According to the feasibility condition, interference from the $f_c(N_A, \bar{d})$ nearest transmitters is eliminated. Hence, the remaining interference is from transmitters $T_i, \forall i = f_n(N, \bar{d}) + 1, f_n(N, \bar{d}) + 2, \ldots$. The outage probability can be expressed as,

$$\epsilon_{ip} = \Pr\left(\frac{s_{0,q}}{\sum_{i=f_n(N,\bar{d})+1}^{\infty} l_{0i}^{-\varsigma} \left\| \mathbf{f}_q^H \mathbf{H}_{0i} \mathbf{V}_i \right\|^2} < \theta \right).$$

Lower Bound of Transmission Capacity

With Markov's inequality,

$$\begin{aligned} \epsilon_{ip} &= \Pr\left(\sum_{i=f_n(N,\bar{d})+1}^{\infty} l_{0i}^{-\varsigma} \left\| \mathbf{f}_q^H \mathbf{H}_{0i} \mathbf{V}_i \right\|^2 > \frac{s_{0,q}}{\theta} \right) \\ &\leq \frac{\theta}{s_{0,q}} \mathbb{E}\left[\sum_{i=f_n(N,\bar{d})+1}^{\infty} l_{0i}^{-\varsigma} \left\| \mathbf{f}_q^H \mathbf{H}_{0i} \mathbf{V}_i \right\|^2 \right]. \end{aligned}$$

Because of the independence between $l_{0i}^{-\varsigma}$ and $\left\| \mathbf{f}_q^H \mathbf{H}_{0i} \mathbf{V}_i \right\|^2$, the upper bound of the outage probability can be obtained,

$$\epsilon_{ip} \leq \frac{\theta}{s_{0,q}} \mathbb{E}[h] \sum_{i=f_n(N,\bar{d})+1} \mathbb{E}\left[l_{0i}^{-\varsigma} \right] \leq \frac{\bar{d}\theta(\lambda\pi)^{\frac{\varsigma}{2}}}{s} \left(\frac{\varsigma}{2} - 1 \right)^{-1} \left(\left\lfloor \frac{N - \bar{d}}{\bar{d}} \right\rfloor - \left\lceil \frac{\varsigma}{2} \right\rceil \right)^{1 - \frac{\varsigma}{2}},$$

A lower bound of the transmission capacity under the imperfect ZF condition can be expressed as,

$$\mathcal{C}_{ip}^{lb} = \frac{Rd(1-\epsilon)}{\pi} \left[\frac{s_{0,q}\epsilon\left(\frac{\varsigma}{2} - 1\right)}{\bar{d}\theta} \left(\left\lfloor \frac{N - \bar{d}}{\bar{d}} \right\rfloor - \left\lceil \frac{\varsigma}{2} \right\rceil \right)^{\frac{\varsigma}{2} - 1} \right]^{\frac{2}{\varsigma}}.$$

Upper Bound of Transmission Capacity

To find a lower bound of the outage probability, interference from the $f_n(N, \bar{d}) + 1$-th nearest transmitter is considered. The corresponding interference power is denoted as $I_{f_n(N,\bar{d})+1}$ which

satisfies the following equation,

$$\mathbb{E}\left[\frac{1}{I_{f_n(N,\bar{d})+1}}\right]=\mathbb{E}\left[\frac{l_{0,f_n(N,\bar{d})+1}^{\varsigma}}{h}\right]=\frac{(\pi\lambda)^{-\frac{\varsigma}{2}}}{(\bar{d}^2-1)}\frac{\Gamma\left(f_n(N,\bar{d})+1+\frac{\varsigma}{2}\right)}{\Gamma(f_n(N,\bar{d})+1)}. \tag{5.9}$$

By the definition of gamma function, the rightmost fraction can be transferred using the following equation:

$$\begin{aligned}\frac{\Gamma\left(f_n(N,\bar{d})+1+\frac{\varsigma}{2}\right)}{\Gamma(f_n(N,\bar{d})+1)}=&\left(f_n(N,\bar{d})+1+\frac{\varsigma}{2}-1\right)\times\\&\left(f_n(N,\bar{d})+1+\frac{\varsigma}{2}-2\right)\times\cdots\times\left(f_n(N,\bar{d})+1+\frac{\varsigma}{2}-\left\lceil\frac{\varsigma}{2}\right\rceil\right),\end{aligned}$$

where there are $\left\lceil\frac{\varsigma}{2}\right\rceil$ terms and $\left(f_n(N,\bar{d})+\frac{\varsigma}{2}\right)$ is the biggest one, therefore,

$$\frac{\Gamma\left(f_n(N,\bar{d})+1+\frac{\varsigma}{2}\right)}{\Gamma(f_n(N,\bar{d})+1)}\leq\left(f_n(N,\bar{d})+\frac{\varsigma}{2}\right)^{\left\lceil\frac{\varsigma}{2}\right\rceil}. \tag{5.10}$$

Taking (5.10) into (5.9), the following inequality can be obtained:

$$\mathbb{E}\left[\frac{1}{I_{f_n(N,\bar{d})+1}}\right]\leq\frac{(\pi\lambda)^{-\frac{\varsigma}{2}}}{(\bar{d}-1)}\left(\left\lfloor\frac{N-\bar{d}}{\bar{d}}\right\rfloor+\frac{\varsigma}{2}\right)^{\left\lceil\frac{\varsigma}{2}\right\rceil}.$$

With Markov's inequality, the following inequality can be obtained, which can be regarded as an upper bound of the success probability:

$$1-\epsilon_{ip}=\Pr\left(\frac{1}{I_{f_n(N,\bar{d})+1}}>\frac{\theta}{s_{0,q}}\right)\leq\frac{s_{0,q}}{\theta}\mathbb{E}\left[\frac{1}{I_{f_n(N,\bar{d})+1}}\right]=\frac{(\pi\lambda)^{-\frac{\varsigma}{2}}s_{0,q}}{\theta(\bar{d}-1)}\left(\left\lfloor\frac{N-\bar{d}}{\bar{d}}\right\rfloor+\frac{\varsigma}{2}\right)^{\left\lceil\frac{\varsigma}{2}\right\rceil}.$$

Hence, an upper bound of the transmission capacity under the imperfect ZF condition can be rewritten as,

$$\mathcal{C}_{ip}^{ub}=R\bar{d}(1-\epsilon)\left[\frac{s_{0,q}\left(\left\lfloor\frac{N-\bar{d}}{\bar{d}}\right\rfloor+\frac{\varsigma}{2}\right)^{\left\lceil\frac{\varsigma}{2}\right\rceil}}{\pi^{\frac{\varsigma}{2}}\theta(\bar{d}-1)}\right]^{\frac{2}{\varsigma}}.$$

5.5 ZF Receiver with IA Beamforming in A Dynamic Network

5.5.1 A More Specific System Model

In this section, both random transmitter distribution and random data traffic are considered. It is assumed that secondary receiver is allowed to join the network only when they can feasibly eliminate all interference in their CSI areas. In this way, $r_z = r_c$ and the interference at a receiver follows two independent stochastic processes. Specifically, the random transmitter distribution forms a 2-dimensional process outside its CSI area and the random data traffic forms a 1-dimensional process inside the CSI area.

It is assumed the transmitters in this case have the same CSI availability as the receivers and perform IA beamforming. Specifically, when a transmitter joins the network, it will try to align its interference to the occupied interference subspaces at the receivers inside its CSI area. That is to say, to a receiver, interference signals are aligned sequentially.

Queueing

It is assumed that the whole network is initialised at the *zero*-th time slot. In other words, there is no active transmitter-receiver pair on the plane before time slot zero. When any transmitter-receiver pair is turned on, its data link will hold until all of the required information is successfully transmitted. The holding times for all of the data links are assumed to be identical and independent exponentially distributed with mean $\frac{1}{\mu_h}$. Its CDF, which is defined as $F_h(x)$, can be written as,

$$F_h(x) = 1 - \exp(-\mu_h x).$$

As stated before, for R_0, the secondary interferer can join the network in the ZF area of R_0 if and only if they can actively eliminate all of its interference to R_0.

Because the interference from transmitters outside the ZF area is treated as shot noise, the interference subspace of R_0 is only affected by the interferers inside its ZF area. In order to better describe the queueing, the indices of the transmitters inside $\mathcal{A}_z$ are reallocated where T_i joins the network at time slot τ_i, where $\tau_i \leq \tau_{i+1}, i \in \mathbb{Z}$. Note that the active transmitters inside $\mathcal{A}_z$ when R_0 joins the network have indices smaller than *zero* and they are called the *dominant transmitters*. The arrival interval between T_i and T_{i+1} is defined as a random variable

with identical and independent exponentially distributed with mean μ_v and the CDF $F_v(x)$ is expressed as,

$$F_v(x) = \Pr\{\tau_{i+1} - \tau_i \leq x\} = 1 - \exp(-\mu_v x), \forall i \in \mathbb{Z}.$$

where μ_v is called the *arrival rate*.

With exponentially distributed holding time and arrival interval, the data traffic from transmitters inside $\mathcal{A}_z$ forms a $M/M/1$ queue, wherein the number of servers in this queueing model can be interpreted as the number of channels. Roughly, the network is divided into two stages: *early stage* and *stable stage*. Based on the queueing theory, in the early stage, the transmitter intensity is increasing and in the stable stage, i.e., time slot t where $t \to \infty$, transmitter will distribute as a PPP with stable intensity [82].

A Toy Example

In Figure 5.1, R_0 is located at the centre with intended transmitter T_0. At time slot 1, T_1 and T_5 join the network and set up their data links. At time slot 2, the intended transmitter-receiver pair (T_0, R_0) joins the network. Receiver R_0 will utilise the ZF method to eliminate interferences from the dominant transmitters T_1 and T_5. The number of interferers that R_0 can eliminate is based on the number of antennas and the number of data streams. In this example, it is assumed that (T_0, R_0) can feasibly join the network. Afterwards, R_0 will define its reference interference subspace. Specifically, it can choose the interference from T_1 or T_5 as the reference interference, and, thus determine the basis of the reference interference subspace. Afterwards, R_0 conveys the basis of the reference interference subspace to all listening transmitters inside $b(R_0, r_z)$, i.e., T_0, T_2, T_3, T_4. Since the second time slot, when new transmitters T_2, T_3, T_4 join the network, they will align their interferences to the reference interference of R_0. Importantly, when T_5 leaves the network, its interference to R_0 disappears, which can reduce the dimension of interference subspace at R_0 if T_5 is not selected as the reference interferer. Consequently, the dimension of the interference subspace at R_0 decreases.

5.5.2 Interference Alignment Implementation

At the time slot when (T_0, R_0) joins the network, R_0 will actively eliminate the interference in $b(R_0, r_z)$. If this can be feasibly accomplished, R_0 will choose the receiving matrix lying in the

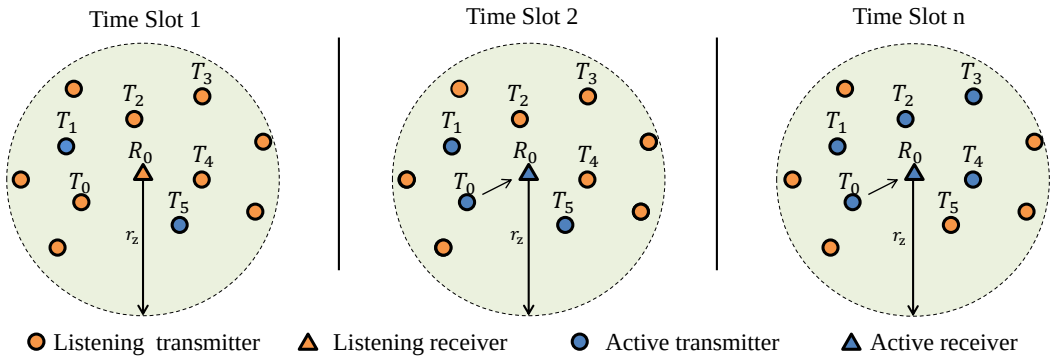


Figure 5.1: Dynamic data traffic around R_0 in n time slots, $n \geq 3$.

null space of all interference from the dominant transmitters, i.e.,

$$\mathbf{U}_0^H \mathbf{H}_{0i} \mathbf{V}_i = \mathbf{0}, \forall T_i \in b(R_0, r_z),\ T_i \text{ is active at time slot } \tau_0.$$

After selecting the reference transmitter T_f among the dominant transmitters, the reference interference subspace is given by $\text{span}[\mathbf{H}_{0f}\mathbf{V}_f]$. All transmitters in $b(R_0, r_z)$ that join the network after τ_0 will align their interference to $\text{span}[\mathbf{H}_{0f}\mathbf{V}_f]$ at R_0, i.e.,

$$\text{span}[\mathbf{H}_{0i}\mathbf{V}_i] = \text{span}[\mathbf{H}_{0f}\mathbf{V}_f], \forall T_i \in b(R_0, r_z), \forall \tau_i > \tau_0,$$

where it is assumed that the number of transmitter antennas is always enough to align the interference. In this way, all transmitters joining $b(R_0, r_z)$ after R_0 will generate no interference to R_0. More importantly, the interference subspace at R_0 will converge from a combination of unaligned interference subspaces to $\text{span}[\mathbf{H}_{0f}\mathbf{V}_f]$. When the dimension of the interference subspace decreases, the R_0 can exploit the unoccupied subspace in the signal space and increase the diversity gain.

Feasibility Condition and Access Probability

In this section, the receiver can cancel all interference inside range $b(R_0, r_z)$. According to the Bezout's theorem, the condition that R_0 can feasibly eliminate all interference is based on the number of equations being no greater than the number of variables. In order to set up the link between T_0 and R_0, $\bar{d}$ data stream must be transmitted, which means R_0 can provide $(N - \bar{d})\bar{d}$ variables. To eliminate interference from one of the dominant transmitters, there are $\bar{d}^2$ equations. Considering that there are $\hat{k}_0$ dominant transmitters in $b(R_0, r_z)$ when at time slot τ_0, the total

number of equations is $\hat{k}_0\bar{d}^2$. Since the dominant transmitters are not changing the beamforming matrices, all of the variables used to cancel the interference are provided by R_0. Therefore the condition that R_0 can feasibly eliminate interference from all of the $\hat{k}_0$ dominant transmitters can be rewritten as,

$$(N-\bar{d})\bar{d} \geq \hat{k}_0\bar{d}^2 \Longleftrightarrow \hat{k}_0 \leq \left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor,$$

where $\lfloor\cdot\rfloor$ is the floor function. Hence, if the proposed transmitter-receiver pair (T_0, R_0) wants to join the network, there should be no more than $\left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor$ dominant transmitters in $b(R_0, r_z)$. Hence, the probability that (T_0, R_0) can successfully join the network is $\Pr\left(\hat{k}_0 \leq \left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor\right)$. Finally, interference occupies $\hat{k}_0\bar{d}$-dimensional subspace at R_0 and the desired signal occupies an independent $\bar{d}$-dimensional subspace.

The probability that a secondary receiver is able to feasibly eliminate interference within their ZF area is called the *access probability*. Note that the feasibility condition means that some transmitter-receiver pairs will not be allowed to join the network. Thus the transmitters actually form a Poisson hole process. As the Poisson hole process can be accurately approximated by PPP [83], the transmitters on the plane are still regarded as Poisson distributed. When calculating the access probability, the expected number of the transmitters in $b(R_0, r_z)$ increases at the early stage of the network and goes up to a constant number when the network is stable. According to [82], the probability of $\hat{k}_0 = k$ at time slot t can be presented as,

$$\Pr(\hat{k}_0 = k \text{ at time slot } t) = \exp\left(-\frac{\mu_v(1-e^{-\mu_h t})}{\mu_h}\right)\frac{\left[\frac{\mu_v(1-e^{-\mu_h t})}{\mu_h}\right]^k}{k!}. \tag{5.11}$$

If the network is in the stable stage when (T_0, R_0) joins the network, (5.11) can be rewritten as,

$$\Pr(\hat{k}_0 = k \text{ at stable stage}) = \lim_{t\to\infty}\Pr(\hat{k}_0 = k \text{ at time slot } t) = e^{-\mu_v/\mu_h}\frac{(\mu_v/\mu_h)^k}{k!}. \tag{5.12}$$

Deduced from (5.11) and (5.12), at time slot t, the transmitters on the plane follow Poisson distribution with the following intensities in the early and stable stages, respectively:

$$\lambda = \frac{\mu_v(1-e^{-\mu_h t})}{\pi\mu_h r_z^2}, \quad \lambda_\infty = \lim_{t\to\infty}\lambda = \frac{\mu_v}{\pi\mu_h r_z^2}.$$

It can be observed that the transmitter intensity λ increases with time until achieving a constant value when the network is stable. The access probability for R_0 when joining the network can be

expressed as,

$$\Pr\left(\hat{k}_0 \leq \left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor \text{ at time slot } t\right) = \frac{\Gamma\left(\lfloor \frac{N-\bar{d}}{\bar{d}} \rfloor + 1, \frac{\mu_v(1-e^{-\mu_h t})}{\mu_h}\right)}{\Gamma\left(\lfloor \frac{N-\bar{d}}{\bar{d}} \rfloor + 1\right)}, \tag{5.13}$$

$$\Pr\left(\hat{k}_0 \leq \left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor \text{ at stable stage}\right) = \frac{\Gamma\left(\lfloor \frac{N-\bar{d}}{\bar{d}} \rfloor + 1, \frac{\mu_v}{\mu_h}\right)}{\Gamma\left(\lfloor \frac{N-\bar{d}}{\bar{d}} \rfloor + 1\right)}. \tag{5.14}$$

From (5.13) and (5.14), μ_v can be adjusted to achieve proper access probability so that the transmitter-receiver pair does not need to wait for a long time. Because it is assumed that all of the transmitters on the plane are homogeneous, the arrival rate in a unit area is defined as $\bar{\mu}_v = \frac{\mu_v}{\pi r_z^2}$. In this way, equations (5.13) and (5.14) can be rewritten as,

$$\Pr\left(\hat{k}_0 \leq \left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor \text{ at time slot } t\right) = \frac{\Gamma\left(\lfloor \frac{N-\bar{d}}{\bar{d}} \rfloor + 1, \frac{\pi r_z^2 \bar{\mu}_v(1-e^{-\mu_h t})}{\mu_h}\right)}{\Gamma\left(\lfloor \frac{N-\bar{d}}{\bar{d}} \rfloor + 1\right)},$$

$$\Pr\left(\hat{k}_0 \leq \left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor \text{ at stable stage}\right) = \frac{\Gamma\left(\lfloor \frac{N-\bar{d}}{\bar{d}} \rfloor + 1, \frac{\pi r_z^2 \bar{\mu}_v}{\mu_h}\right)}{\Gamma\left(\lfloor \frac{N-\bar{d}}{\bar{d}} \rfloor + 1\right)}.$$

where the access probability is actually a function of the range r_z if $\bar{\mu}_v$ and μ_h are constant. On the one hand, if r_z is too large, the access probability will be low. On the other hand, if r_z is too small, the remaining interference will be large. Because a secondary transmitter-receiver pair can only join the network if the interference in the ZF area of the receiver can be eliminated, the access probability can be interpreted as a delay of the transmission between the transmitter-receiver pair [84].

Interference Subspace Evolution

As stated before, all interferences from transmitters in $\mathcal{A}_z$ are aligned to T_f since (T_0, R_0) successfully joins the network. Consequently, the dimension of interference subspace at R_0 is shrinking. At time slot t, the dimension of interference-free subspace at R_0 is denoted as $\bar{d}_0(t)$.

The number of dominant transmitters in $b(R_0, r_z)$ is assumed to be $\hat{k}_0$ when R_0 joins the network and the expected remaining holding time for all of the $\hat{k}_0$ transmitters is assumed to be $\frac{1}{c\mu_h}$, where $c \geq 1$. It will be shown in the simulation that having $c = \frac{\hat{k}_0}{2}$ accurately fits the real situation. Therefore, the probability that one of the $\hat{k}_0$ transmitters has already left the network before time slot t is given by $1 - e^{-c\mu_h(t-\tau_0)}$. All transmitters in $\mathcal{A}_z$ that join after

R_0 align their interferences to the reference interference subspace at R_0, therefore the dimension of the interference-free subspace increases by $\bar{d}$ whenever the transmission from one dominant transmitter $\{T_0|i < 0, i \neq f\}$ finishes. This process continues until the interference subspace converges to the $\bar{d}$-dimensional reference subspace and $\bar{d}_0(t) = N - \bar{d}$. The probability that n of the $\hat{k}_0$ dominant transmitters have left the network at time slot t can be written as,

$$\binom{\hat{k}_0 - 1}{n} \left(1 - e^{-c\mu_h(t-\tau_0)}\right)^n \left(e^{-c\mu_h(t-\tau_0)}\right)^{\hat{k}_0 - n - 1}.$$

where $\binom{\hat{k}_0 - 1}{n}$ is the binomial coefficient. This binomial distribution has an expected value of $(\hat{k}_0 - 1)(1 - e^{-c\mu_h(t-\tau_0)})$, i.e., the expected number of dominant transmitters finished transmitting at time slot t can be obtained as $\hat{k}_0 - 1 - (\hat{k}_0 - 1)(1 - e^{-c\mu_h(t-\tau_0)}) = (\hat{k}_0 - 1)e^{-c\mu_h(t-\tau_0)}$. Therefore, the expected number of remaining dominant transmitters at time slot t can be expressed as $1 + (\hat{k}_0 - 1)e^{-c\mu_h(t-\tau_0)}$. Hence, the expected dimension of interference-free subspace at R_0 at time slot t can be expressed as,

$$\bar{d}_0(t) = N - \bar{d}\left(1 + (\hat{k}_0 - 1)e^{-c\mu_h(t-\tau_0)}\right). \tag{5.15}$$

In order to approximate the real situation, $\hat{k}_0$ is set to be the the average number of transmitters in $b(R_0, r_z)$ on condition that it is no more than $\left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor$. Hence,

$$\mathbb{E}[\hat{k}_0] = \frac{\Gamma\left(\lfloor \frac{N-\bar{d}}{\bar{d}} \rfloor + 1\right)}{\Gamma\left(\lfloor \frac{N-\bar{d}}{\bar{d}} \rfloor + 1, \frac{\pi r_z^2 \bar{\mu}_v (1 - e^{-\mu_h t})}{\mu_h}\right)} \sum_{k=0}^{\left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor} \exp\left(-\frac{\mu_v(1 - e^{-\mu_h t})}{\mu_h}\right) \frac{\left(\frac{\mu_v(1-e^{-\mu_h t})}{\mu_h}\right)^k}{(k-1)!}.$$

If the network is in stable stage, i.e., $t \to \infty$, then,

$$\mathbb{E}[\hat{k}_0] = \frac{\Gamma\left(\lfloor \frac{N-\bar{d}}{\bar{d}} \rfloor + 1\right)}{\Gamma\left(\lfloor \frac{N-\bar{d}}{\bar{d}} \rfloor + 1, \frac{\pi r_z^2 \bar{\mu}_v}{\mu_h}\right)} \sum_{k=0}^{\left\lfloor \frac{N-\bar{d}}{\bar{d}} \right\rfloor} \exp\left(-\frac{\mu_v}{\mu_h}\right) \frac{\left(\frac{\mu_v}{\mu_h}\right)^k}{(k-1)!}.$$

As predicted, the interference subspace will tend to be $\bar{d}$ as time goes on if all coming interferences are aligned, i.e., the dimension of the interference-free subspace tends to be the following equation:

$$\lim_{t \to \infty} \bar{d}_0(t) = N - \bar{d}.$$

5.5.3 Outage Probability and Transmission Capacity

In the case of Rayleigh fading, the Laplace transform of interference I_q can be expressed as,

$$\mathcal{L}_{I_q}(s) = \mathbb{E}\left[\exp\left(-hs \sum_{T_i \in \Phi_T / b(R_0, r_z)} l_{0i}^{-\varsigma}\right)\right].$$

Because the channel matrices between interferers and the receiver R_0 have i.i.d complex Gaussian entries, the fading coefficient h is the sum of $2\bar{d}$ independent normal distributions, or equivalently, chi-squared distribution with $2\bar{d}$ DoF. The probability density function (PDF) is given by,

$$f_{\chi^2}(x, 2\bar{d}) = \frac{x^{\bar{d}-1} e^{-x/2}}{2^{\bar{d}} \Gamma(\bar{d})}. \tag{5.16}$$

where $f_{\chi^2}(x, 2\bar{d})$ is the PDF of chi-squared distribution with $2\bar{d}$ DoF. Because all of the transmitters are Poisson distributed on the plane with intensity λ, the radius transmitter intensity is given by,

$$f_\Omega(x) = \frac{2x}{r_\infty^2 - r_z^2}, \tag{5.17}$$

where r_∞ is the supposed distance between the receiver R_0 and the furthest transmitter on the plane. Considering the Poisson distribution with intensity λ and the number of transmitters outside $b(R_0, r_z)$ is k, the success probability is given by,

$$\mathcal{L}_{I_q}(s) = \sum_{k=0}^{\infty} \frac{\exp(-\pi\lambda(r_\infty^2 - r_z^2))(\pi\lambda(r_\infty^2 - r_z^2))^k \mathcal{Q}_6}{k!},$$

where $\mathcal{Q}_6$ is the success probability if there are exactly k transmitters outside $\mathcal{A}_z$, i.e.,

$$\mathcal{Q}_6 = \underbrace{\int_{r_z}^{r_\infty} \cdots \int_{r_z}^{r_\infty}}_{k} \int_0^\infty \exp\left(-\frac{\theta h l_{00}^{\varsigma}}{2} \sum_{T_0 \in \Phi_T / b(R_0, r_z)} \omega^{-\varsigma}\right) f_{\chi^2}(x, 2\bar{d}) f_\Omega(\boldsymbol{\omega}) dx d\boldsymbol{\omega},$$

$\boldsymbol{\omega}$ here is a k-dimensional vector representing the distances between transmitters to R_0.

Theorem 8 *Under Rayleigh fading, the Laplace transform of the interference I_q with a fixed ZF area can be written as,*

$$\mathcal{L}_{I_q}(s) = \exp\left(\frac{2^{\bar{d}} r_z^2 - (\theta l_{00}^{\varsigma} r_z^{-\varsigma}/2 + 1/2)^{-\bar{d}} r_z^2 - s\varsigma\bar{d}\mathcal{Q}_7/2}{(\pi\lambda)^{-1} 2^{\bar{d}}}\right), \tag{5.18}$$

where,

$$\mathcal{Q}_7 = \frac{2^{1+\bar{d}} r_z^{2-\varsigma} {}_2F_1\left[\frac{\varsigma-2}{\varsigma}, 1+\bar{d}, 2-\frac{2}{\varsigma}, -sr_z^{-\varsigma}\right]}{\varsigma-2},$$

and ${}_2F_1[\cdot]$ *is the hypergeometric function.*

Proof: See Appendix C

Note that Laplace transform of interference actually represents the success probability when the signal power is exponentially distributed.

Success Probability

Recall that the dimension of interference-free subspace at R_0 has been characterised by (5.15). Hence, the efficient SRDoF that can be utilised to decode any data stream is $\hat{d}_0(t) = \bar{d}_0(t) - \bar{d} + 1 = N - 2\bar{d} - \bar{d}(\hat{k}_0 - 1)e^{-c\mu_h(t-\tau_0)} + 1$. As discussed in the system model (5.1), the desired signal power becomes chi-squared distributed with $2\hat{d}_0(t)$ DoF whose complementary CDF is $\tilde{F}_\Gamma(x) = \frac{\Gamma(\hat{d}_0(t),x)}{\Gamma(\hat{d}_0(t))}$. In this way, the success probability can be expressed as,

$$p_s = \sum_{k=0}^{\lfloor \hat{d}_0(t) \rfloor - 1} \frac{(-\theta l_{00}^{\varsigma})^k}{k!} \frac{\mathrm{d}^k}{\mathrm{d}s^k} \mathcal{L}_{I_q}(s)\Big|_{s=\theta l_{00}^{\varsigma}}, \tag{5.19}$$

which is from the derivative property of the Laplace transform and Theorem 8.

One important observation from this expression is that the success probability increases with $\hat{d}_0(t)$ because all of the terms inside the summation are positive. Intuitively, with larger μ_h, i.e., smaller expected holding time, the success probability increases more quickly.

Transmission Capacity

Practically, the transmitters may not always be able to perfectly align their interference to all of the receivers in their ZF area when joining the network. Given this, the interference may randomly occupy the extra subspace that is not occupied by the desired signal or existing interference. Hence, the interference subspace at R_0 may not be shrinking towards $\bar{d}$ and the extra interference subspace can not be exploited to generate diversity when decoding the desired symbols. Because there are $\bar{d}$ SRDoF in total to decode the $\bar{d}$ independent data streams, there is only one SRDoF available to decode one private symbol, i.e., the signal power becomes exponentially distributed. Therefore, the Laplace transform of the interference I_q at the point $s = \theta l_{00}^{\varsigma}$ becomes the success

probability of the q-th data stream at R_0, by which the following expression can be obtained:

$$\lambda^{\epsilon} = \frac{2^{\bar{d}} \ln(1-\epsilon)}{\pi 2^{\bar{d}} r_z^2 - \pi(\theta l_{00}^{\varsigma} r_z^{-\varsigma}/2 + 1/2)^{-\bar{d}} r_z^2 - \pi\varsigma\theta\bar{d} l_{00}^{\varsigma} \mathcal{Q}_7/2}. \tag{5.20}$$

Considering the definition of transmission capacity in (2.6), it can be given by,

$$\mathcal{C} = \frac{R\bar{d}2^{\bar{d}}(1-\epsilon)\ln(1-\epsilon)}{\pi 2^{\bar{d}} r_z^2 - \pi(\theta l_{00}^{\varsigma} r_z^{-\varsigma}/2 + 1/2)^{-\bar{d}} r_z^2 - \pi\varsigma\theta\bar{d} l_{00}^{\varsigma} \mathcal{Q}_7/2}. \tag{5.21}$$

5.6 Simulation Results

In this section, the numerical simulations of how the results derived in this chapter are applied are presented. The specific steps for Monte Carlo simulations (outage probability) are given below. Note that the Monte Carlo simulation for SINR has similar steps.

1. Define the transmitter intensity and the edge length of the panel L. The total number of transmitters is a Poisson random variable with mean λL^2. In all of the following simulations, at least 200 transmitters are expected to be included on the panel.

2. Generate coordinates for the transmitters the x-axis and y-axis coordinates of which are uniformly distribution between $-\frac{L}{2}$ to $\frac{L}{2}$.

3. Assume the receiver is at the centre. Remove the transmitters the distances to the centre of which are less than r_z.

4. Assign each remaining transmitter with an i.i.d. chi-squared fading coefficient. The total interference power is the sum of the products of the fading coefficients and the path-loss effect.

5. Assume the desired transmitter is at distance l_{00} from the centre. The signal power can be expressed as a product of $l_{00}^{-\alpha}$ and an exponentially or a chi-squared distributed random variable.

6. Repeat the above steps 3000 times at least. In each repeat, if the SIR is smaller than the unspecified threshold, add the number of outage events by *one*. Finally, dividing the total number of outage events by the total number of repeats will give the outage probability.

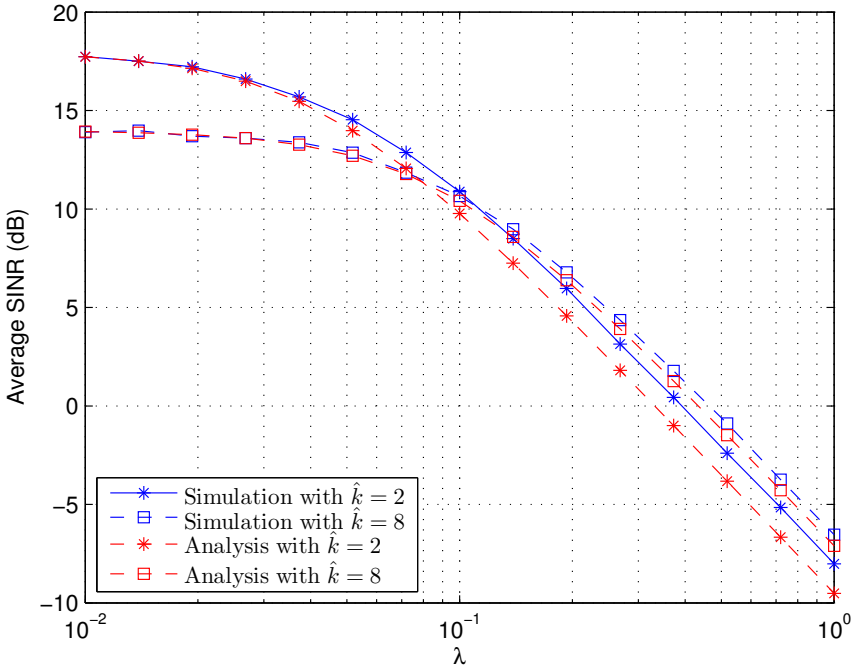


Figure 5.2: Average signal-to-noise ratio of a ZF receiver with various coverage factors.

Figure 5.2 depicts the average SINR derived in (5.3) versus transmitter intensity where $\varsigma = 4$, $N = 12$ and $l_{00} = 2$. The SNR at the transmitter side is set to be 20 dB, i.e., $n_0 = 0.01$. From this figure, it can be seen that (5.3) almost accurately models the average SINR. Note that a larger coverage factor $\hat{k}$ means more interference is cancelled but the SRDoF to decode the desired symbol, which is equal to $N - \hat{k}$, decreases. Interestingly, the average SINR achieved with larger $\hat{k}$ outperforms that with smaller $\hat{k}$ when the transmitter intensity is large. This is because larger transmitter intensity means the interference power generated by the nearest transmitters becomes so dominant that severe SINR degradation occurs if uncancelled. On the other side, when the transmitter intensity decreases, SINR with smaller $\hat{k}$ becomes the choice. The reason behind this is that low transmitter intensity means interference power from the nearest transmitters becomes comparable to the additive white noise. Hence, there is no need to cancel their interference, and allocating more SRDoF to decode the desired symbol becomes efficient to increase the SINR. Conclusively speaking, in the high transmitter intensity region, the receiver should allocate more SRDoF to cancel the interference.

Figure 5.3 shows the access probability versus the holding time factor μ_h where $N = 4$, $\mu_v = 1$ and transmitters are sending $\bar{d} = 1, 2$ data streams. Early stage network and stable stage network are depicted where the corresponding expressions are (5.13) and (5.14), respectively. For the early stage lines, $t = 8$ is considered. In the stable stage, the transmitter intensity is given by

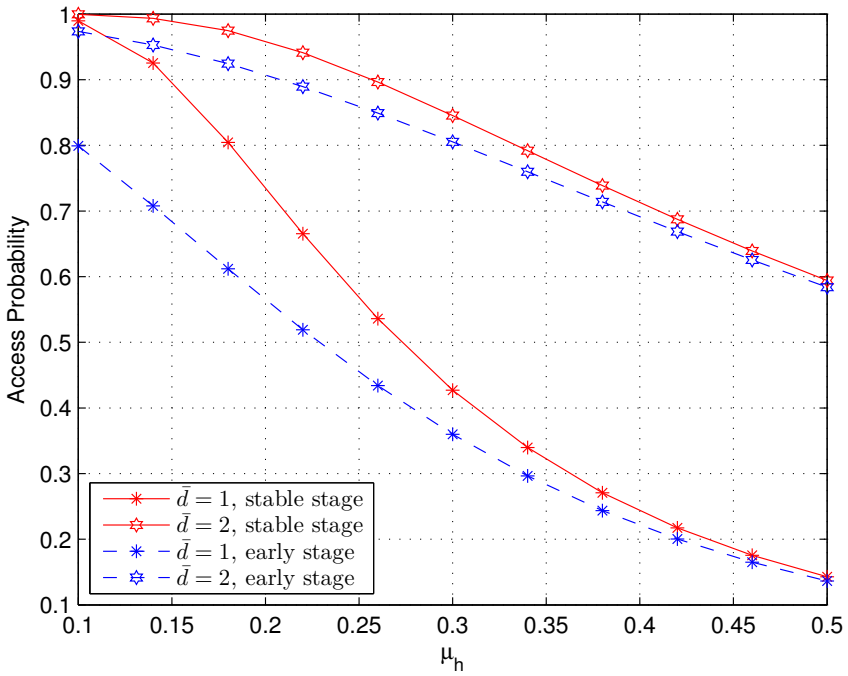


Figure 5.3: Access probability versus μ_h in early stage and stable stage.

$\frac{\mu_v}{\pi\mu_h r_z^2}$. Thus, increasing the holding time is equivalent to increasing the transmitter intensity and thus lowers the access probability. It can be seen that the access probabilities in the early stage approach the access probabilities in the stable stage more quickly when μ_h increases. This is because the transmitter intensity in the early stage approaches that in the stable stage at the rate of $1-e^{-\mu_h t}$. One important observation from this figure is that the gap between access probabilities with $\bar{d}=1$ and $\bar{d}=2$ increases with μ_h. It indicates that with large transmitter intensity, more access probability is sacrificed to increase the number of data streams in the network.

Recall that in order to derive the dimension of interference-free subspace in (5.15), it was assumed that all the $\hat{k}_0$ dominant transmitters have the same expected remaining holding time $\frac{1}{c\mu_h}$, where $c \geq 1$. Meanwhile, having $c = \frac{\hat{k}_0}{2}$ can approximate the real world situation properly was also stated. This statement for a stable stage network is validated in Figure 5.4 where (5.15) with different μ_h as well as the corresponding simulation results are shown. In order to run fair Monte Carlo simulations, each transmitter is assigned an independent remaining holding time, which is exponentially distributed with the mean value being randomly selected between $\frac{1}{\mu_h}$ and $\frac{1}{\mu_h \hat{k}_0}$. It can be seen that our assumption brings tractability to our result without scarifying significant accuracy.

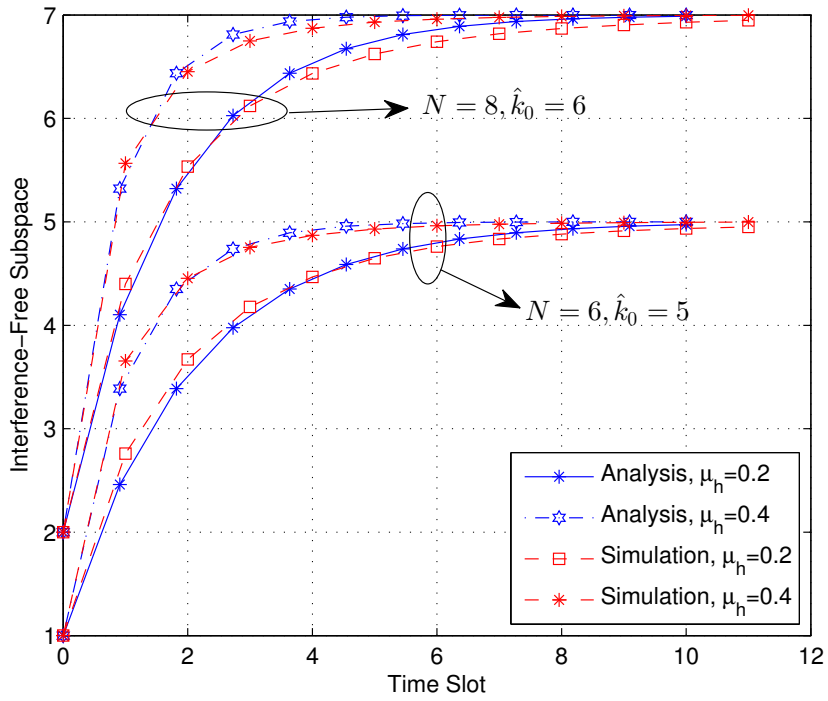


Figure 5.4: Analysis and simulation of dimension of the interference-free subspace.

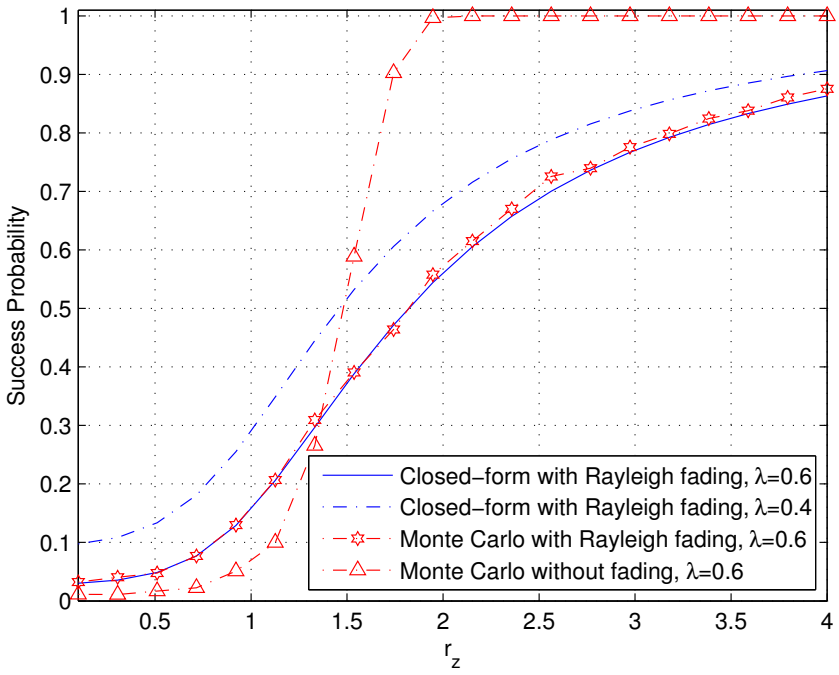


Figure 5.5: Analysis and simulation of the success probability versus r_z.

Figure 5.5 investigates the Laplace transform of the interference versus radius of the ZF area r_z, where the analysis lines from (5.18) are depicted as well as the Monte Carlo simulations with and

without fading. Note that the line without fading is merely a theoretical bound when the fading becomes extremely slow. Recall that this Laplace transform represents the success probability at any data stream of R_0 if *one* SRDoF is utilised to decode a private symbol. In the simulation, the distance between the intended transmitter-receiver pair is $l_{00} = 0.5$, the number of data streams is $\bar{d} = 2$, the pre-specified SIR threshold is $\theta = 10$ and the path-loss exponent is $\varsigma = 4$. Intuitively, the success probability increases with larger r_z and smaller λ because both of them weaken the interference power. One interesting observation is that the line without fading underperforms the line with fading when the ZF area is small. This is because with a given l_{00}, smaller r_z makes the interference power larger compared to the desired signal power. Therefore, the interference power is reduced more significantly by the fading effect compared to the desired signal power. Conversely, when the interference power is reduced below a certain level, e.g., when $r_z = 1.4$ in this figure, the fading effect reduces the desired signal power more significantly and thus, the success probability achievable without fading becomes higher. As a conclusion, fading effect is beneficial when the distance between the transmitter-receiver pair is large. In other words, for a fixed ZF area, in order to make use of the fading effect, a transmitter should select a distant intended receiver.

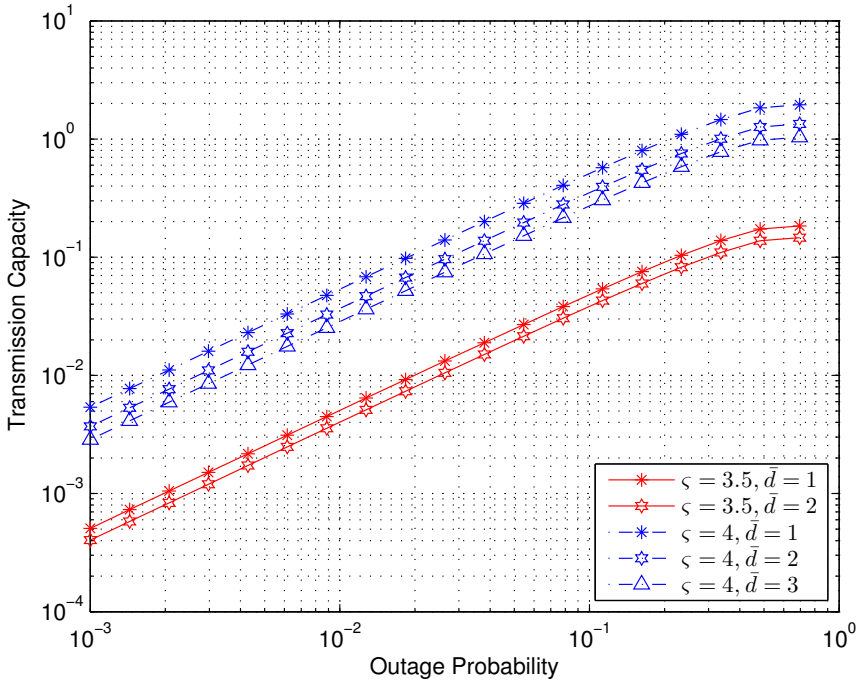


Figure 5.6: Analysis and simulation of the transmission capacity versus outage probability.

Finally, Figure 5.6 demonstrates the closed-form expression of the transmission capacity (5.21) versus the outage probability when $r_z = 5$, $\theta = 10$ and $l_{00} = 1$. Various numbers of data

streams $\bar{d} = 1, 2, 3$ are assumed. Intuitively, higher transmission capacity is achieved with larger ς as interference decreases faster than the desired signal. Meanwhile, the transmission capacity is highly sensitive to the path-loss exponent, i.e., increasing ς by 0.5 may cause 10 times transmission capacity enhancement. One interesting observation is that the transmission capacity is lower with a larger number of data streams $\bar{d}$. This is because with larger number of data streams, the SIR at each data stream of R_0 decreases which results in lower maximum transmitter intensity that guarantees the outage probability. As a results, the transmission capacity enhancement brought by the multiplexing gain is offset by the decreasing transmitter intensity. Conclusively speaking, with a fixed ZF area radius, it is optimal to transmit one data stream. This conclusion is consistent with that proposed in [9].

5.7 Summary

In this chapter, the performance of the ZF receiver in homogeneous Poisson distributed wireless networks was analysed in which all of the transmitters and receivers were equipped with multiple antennas. Note that the performance of the analysed receiver represented the average performance of all the receivers. Average SINR, success probability and transmission capacity were selected as the metrics wherein transmission capacity is known as a tailor-made metric for large-scale randomly distributed wireless networks.

This chapter began with a simple ZF receiver in which the interference from the transmitters within its CSI area was eliminated. The radius of the CSI area determined the number of cancelled interferers, i.e., the coverage factor. By designing the radius of CSI area, the receiver exploited part of the SRDoF to cancel interference and the remaining SRDoF to decode its desired symbol. The closed-form expressions of the average SINR as well as the bounds of the transmission capacity were derived. As a result, it was shown that, in the case of low transmitter intensity, more SRDoF should be allocated to decode its desired symbol.

The impact of the feasibility condition of the secondary ZF receiver to the performance was analysed in the presence of finite CSI area. Transmission capacity was derived as the weighted mean of the two instances according to whether the ZF receiver can feasibly eliminate all interference in the CSI area. Afterwards, a completely dynamic and randomly distributed wireless network was considered, in which the joining and leaving process of the transmitters were modelled by queueing theory. In order to characterise the interference, three jobs were

accomplished. Firstly, the feasibility condition of the ZF receiver was studied and it determined the probability that the receiver was allowed to join an early or stable stage network. Next, the evolving success probability at R_0 was characterised using the queueing model, in which it can be seen that the success probability increased more quickly when the expected holding time was short. Finally, the closed-form expression of the transmission capacity was derived. Two important conclusions were: 1) In a rich scattering environment, a transmitter should select a distant intended receiver in order to make use of the fading effect; 2) With a fixed CSI area, it is optimal to transmit only one data stream, which is beneficial not only to the area spectrum efficiency, but also to the access probability.

Chapter 6

Throughput for A Generic Heterogeneous Wireless Network Using Interference Alignment

6.1 Introduction

Stochastic geometry-based approaches have been extensively exploited to model and analyse large-scale randomly distributed wireless networks [10, 23, 71, 73]. However, most of the analysis was restricted to simple random arrangement of nodes such as PPP and single-antenna setting for tractability reasons. One extension of the aforementioned stochastic geometry-based analysis can be found in [85], where the authors modelled the interference in clustered networks. The probability generating function (PGF) with respect to the Palm measure was studied and thereby, the outage probability of a single link and the transmission capacity of the network were derived. With this PGF, the performance of different interference cancellation techniques can be evaluated in clustered wireless networks.

In large-scale clustered wireless networks, IA is traditionally implemented independently in clusters and all of the inter-cluster transmitters will cause non-aligned interference at the receivers. The major reason behind is that the number of transceiver pairs that can cooperate through MIMO is limited by the feasibility condition [35]. For example, based on the PGF in [85], the authors in [86] and [87] studied the performance of PCP wireless networks in which perfect intra-cluster CSI was available and therefore, the intra-cluster interference was completely eliminated using the IA approach. Similarly, the authors in [88] studied the transmission capacity of a PCP wireless network in which IA approach was implemented to handle the intra-cluster interference. Unlike [86] and [87], imperfect CSI was considered in [88], the quality of which is modeled by a distance-dependent function. Although the quality of CSI was distance-dependent, implementation of the IA approach was in the traditional way, i.e., fully implement IA in each cluster. Because this traditional method to implement IA as in [86–88] may not be optimal to achieve the highest spectrum efficiency in the presence of the path-loss effect, one objective of this chapter is to

implement the IA approach in a more effective way, i.e., distance-dependently. Moreover, none of [86–88] takes the cross-tier interference into consideration, i.e., the IA implementation is only suitable for single-tier homogeneous networks. In recent years, heterogeneity has become a key feature of wireless networks and, therefore, another objective of this chapter is to investigate the performance of IA beamforming in a heterogeneous wireless network.

Two-tier heterogeneous wireless networks have drawn a considerable amount of attention in the past couple of years especially for single-antenna cellular networks. The authors in [89] considered a heterogeneous cellular network in which the macro BSs followed PPP and the femto BSs followed PPP or PCP. The trade-off between the performance of the subscriber and non-subscriber femto BS users was studied based on multi-carrier techniques. The authors in [90] also considered a heterogeneous cellular network with two PPP tiers of BSs, namely the macro BS tier and the femto BS tier, and one Neyman-Scott process tier of the femto BS users. The optimal spectrum allocation policy was derived and the aggregated network throughput is maximised with constraint on the quality of service. There are two major differences between this chapter and [89, 90]. On the one hand, in their work, the interferer distribution after spectrum allocation was considered as a simple thinning process of the original distribution. In this chapter, interference is cancelled in an ordered manner; thus, a correlation between the cancelled interference and the remaining interference exists. On the other hand, single-antenna setting was considered in [89, 90], while multiple-antenna setting is considered in this chapter. With the introduction of MIMO system in Section 2.2, it can be seen that the multiple-antenna setting is more practically appealing and the accompanying interference environment is more complex. Another related work is [44], the objective of which was to propose a tractable Gamma approximation of the interference in heterogeneous wireless networks in which two tiers of PPP transmitters were assumed outside a guard region of each receiver.

6.1.1 Contributions

In this chapter, a generic multiple-antenna, 2-tier, randomly distributed, heterogeneous wireless network was considered. Transmitters in tier-A are Poisson distributed with high coverage range, while transmitters in tier-B are Poisson clustered distributed with low coverage range. The throughput at tier-B receivers in each cluster, is derived and maximised. Each tier-B receiver has four kinds of interferences: the nearest cross-tier interference, the remaining cross-tier interference, the inter-cluster interference and the intra-cluster interference, where the intra-cluster

interference is always fully eliminated.

Towards deriving the throughput, the nearest cross-tier interference is mitigated at the tier-B receivers using a distance-dependent IA approach. Specifically, transmitters in tier-B partially align their intra-cluster interference to the nearest cross-tier interference overheard by the receivers in the same cluster in order to increase the multiplexing gain for their own data links. When the distance from the nearest cross-tier interferer to the tier-B receiver is large, the nearest cross-tier interference tends to be treated as shot noise and when the distance is small, it tends to be fully cancelled. With this distance information, the throughput is maximised by compromising between enhancing the SIR at each data stream and increasing the number of data streams. Note that the gain of the accurate distance information compared to the statistical distance information is studied and validated by the simulation results. Moreover, unlike [86–88] in which the feasibility condition of IA was satisfied by default, the feasibility condition of IA is studied in this chapter and explicitly influences the throughput.

Apart from the nearest cross-tier interference and the intra-cluster interference, stochastic geometry-based approaches are utilised to model the remaining interference, which consists of the remaining cross-tier interference and the inter-cluster interference. On the one hand, the interference from the remaining cross-tier transmitters is marginalised so that it affects the outage probability independently. Then the near-closed-form expression of the success probability affected by the marginalised remaining cross-tier interference is derived. On the other hand, outage probability affected by the inter-cluster interference is derived as well as its closed-form approximation. All of the derived expressions for the success probabilities are validated by Monte Carlo simulations.

6.1.2 Organisations

The reminder of this chapter is organised as follows. Section 6.2 introduces the system model, including the distribution of the transmitters, CSI availability, interference categorisation and performance metrics. Section 6.3 presents the implementation of the distance-dependent IA approach to mitigate the nearest cross-tier interference. Expressions of the success probabilities affected by the remaining cross-tier interference and the inter-cluster interference are derived in Section 6.4, which is followed by the simulation results in Section 6.5. Finally, summaries can be found in Section 6.6.

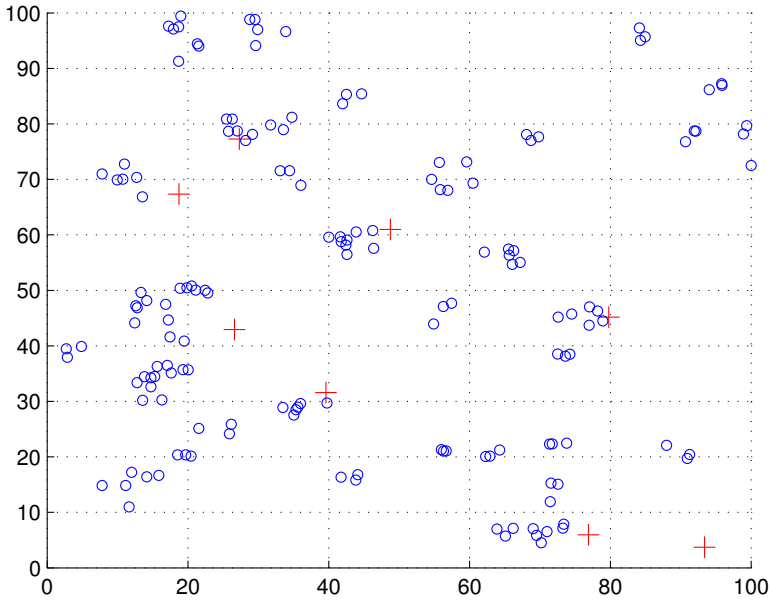


Figure 6.1: The distribution of transmitters in tier-A and tier-B.

6.2 System Model

This chapter considers 2-tiers of transmitters on an infinite plane, namely tier-A and tier-B. The tier-A transmitter set, denoted as $\Phi_A = \{L_i | i \in \mathbb{N}\}$, follows PPP with intensity λ_A and the tier-B transmitter set, denoted as $\Phi_B = \{T_i | i \in \mathbb{N}\}$, follows the Matérn process on the plane with the intensity of the parent process being λ_B. In each cluster of Φ_B, there are K transmitters uniformly distributed in a ball centred at the parent point with radius r_a. For T_i in tier-B, there is an intended receiver, denoted as R_i, distance l_{ii} away, where l_{ii} is a constant value. In order to simplify the following descriptions, the term *tier-B receiver* is used to denote the intended receiver of the tier-B transmitter. The linear density function of the transmitters in each tier-B cluster is given by,

$$f_B(x) = \begin{cases} \frac{2\|x\|}{r_a^2}, & \|x\| \leq r_a \\ \\ 0, & \|x\| > r_a \end{cases},$$

where it is assumed that r_a is small enough so that all the transmitters and receivers in a cluster have the same nearest cross-tier transmitter. The distribution of the tier-A and tier-B transmitters is shown in Figure 6.1, where the red symbols stand for tier-A transmitters and blue symbols represent tier-B transmitters. In this figure, $\lambda_A = 0.005$, $\lambda_B = 0.001$, $K = 3$ and $r_a = 5$.

Note that the transmitter and receiver distribution may represent various tiered wireless network structures such as MANETs and the LTE-A cellular network. In the case of LTE-A cellular networks, which consist of macro BS, pico BS, femto BS and relays, each layer of devices has unique transmit power constraint, distribution, number of antennas, etc. In terms of the proposed two-layer distribution, tier-A transmitters can be thought of as macro BSs while tier-B transmitters can be considered as femto BSs and tier-B receivers represent the user terminals. Macro BSs follow PPP as assumed in [91] and [92] and femto BSs follow PCP as discussed in Section 2.3.7. Macro BSs typically transmit with a high power level (5 W - 40 W) and cover an area with a radius of up to 2 km, while femto BSs typically have a low transmit power level (100 mW - 2 W) and cover an area of up to tens of meters [93]. Therefore, in the presence of path-loss effect, unmitigated interference from tier-A transmitters, especially the nearest one, and the intra-cluster interference may cause poor tier-B system performance.

Each transmitter in tier-A is assumed to transmit $\bar{d}_A$ data streams with N_A antennas and each transmitter in tier-B is assumed to transmit $\bar{d}_B$ data streams with N_B antennas. Tier-B receivers are also equipped with N_B antennas. The signal model for the typical receiver R_0 was described in detail, which is in the Voronoi cell of L_0 and its intended transmitter is in ϕ_0, where ϕ_i is defined as the i-th cluster in tier-B. The signal received at R_0 can be written as,

$$\begin{aligned}\mathbf{y}_0 =& l_{00}^{-\varsigma/2}\mathbf{H}_{00}\mathbf{V}_0\mathbf{s}_0 + \underbrace{\sum_{L_i\in\Phi_A} l_{0L_i}^{-\varsigma/2}\mathbf{H}_{0L_i}\mathbf{V}_{L_i}\mathbf{s}_{L_i}}_{\text{Interference from tier-}A} + \underbrace{\sum_{T_i\in\phi_0} l_{0i}^{-\varsigma/2}\mathbf{H}_{0i}\mathbf{V}_i\mathbf{s}_i}_{\text{Intra-cluster interference}} \\ &+ \underbrace{\sum_{T_i\in\Phi_B/\phi_0} l_{0i}^{-\varsigma/2}\mathbf{H}_{0i}\mathbf{V}_i\mathbf{s}_i}_{\text{Inter-cluster interference}} + \mathbf{n}_i.\end{aligned}$$

Recall that ς denotes the path-loss exponent and

- l_{0L_i} and l_{0i} are the distances between tier-A transmitter L_i and tier-B transmitter T_i to R_0, respectively;
- $\mathbf{H}_{0L_i} \in \mathbb{C}^{N_B\times N_A}$ and $\mathbf{H}_{0i} \in \mathbb{C}^{N_B\times N_B}$ stand for the channels from tier-A transmitters L_i and tier-B transmitter T_i to R_0, respectively. All the entries of channel matrices are i.i.d. $\mathcal{CN}(0,1)$;
- $\mathbf{V}_{L_i} \in \mathbb{C}^{N_B\times\bar{d}_B}$ and $\mathbf{V}_i \in \mathbb{C}^{N_A\times\bar{d}_A}$ are the beamforming matrices with unit-norm columns designed by tier-A transmitter L_i and tier-B transmitter T_i, respectively;
- $\mathbf{s}_i \in \mathbb{C}^{\bar{d}_B\times 1}$ and $\mathbf{s}_{L_i} \in \mathbb{C}^{\bar{d}_A\times 1}$ are symbol vectors transmitted by tier-A transmitter L_i and

tier-B transmitter T_i, respectively;

- $\mathbf{n}_i \in \mathbb{C}^{N_B \times 1}$ is the additive white noise.

It is assumed that all of the symbols in each symbol vector from tier-A transmitters have power P_A and all the symbols in each symbol vector from tier-B transmitters have power P_B. For simplicity, $P_B = 1$ W is considered without loss of generality because P_A can be any positive number and noise is ignored.

The transmitter has access to a high quality of CSI when the receiver is nearby according to [88]. Hence, perfect CSI is shared between transmitter-receiver pairs in a tier-B cluster because the cluster size is small. As the interference from the nearest transmitter is dominant [23] and each BS in tier-A only serves receivers in its Voronoi cell [89], it is assumed that the tier-B transmitters and receivers only have access to CSI of the nearest tier-A transmitter. Moreover, as each tier-A transmitter covers a large area, it can be reasonably assumed that the CSI of the nearest tier-A transmitter at tier-B transmitters and receivers is subject to estimation error. The relationship between the real channel and the corresponding estimate at R_0 can be written as,

$$\mathbf{H}_{0L_0} = \sqrt{1-\varepsilon^2}\hat{\mathbf{H}}_{0L_0} + \varepsilon\tilde{\mathbf{H}}_{0L_0},$$

where $\varepsilon \in [0,1]$ is the standard error of the channel estimation error. Symbol $\hat{\mathbf{H}}_{0L_0}$ is the estimated channel matrix and $\tilde{\mathbf{H}}_{0L_0}$ is the normalised channel estimation error matrix. The Voronoi tessellation of the plane and ϕ_0 are shown in Figure 6.2.

The desired signal is $\bar{d}_B$-dimensional and the interference signals are confined in the remaining $(N_B - \bar{d}_B)$-dimensional subspace at the tier-B receivers [88]. Because all of the intra-cluster interference is eliminated, therefore, the post-processing signal at R_0 can be expressed as,

$$\hat{\mathbf{y}}_0 = l_{00}^{-\varsigma}\mathbf{U}_0^H\mathbf{H}_{00}\mathbf{V}_i\mathbf{s}_0 + \mathbf{U}_0^H\sum_{L_i\in\Phi_A} l_{0L_i}^{-\varsigma/2}\mathbf{H}_{0L_i}\mathbf{V}_{L_i}\mathbf{s}_{L_i} + \mathbf{U}_0^H\sum_{T_i\in\Phi_B/\phi_0} l_{0i}^{-\varsigma/2}\mathbf{H}_{0i}\mathbf{V}_i\mathbf{s}_i,$$

where $\mathbf{U}_0 \in \mathbb{C}^{N_B\times\bar{d}_B}$ is a valid ZF equaliser for IA at R_0 and the intra-cluster interference is eliminated with the help of perfect CSIT. At the q-th data stream of R_0, the desired signal power, denoted as $s_{0,q}$, the cross-tier interference power, denoted as I_c, and the inter-cluster interference power, denoted as I_{inter}, are given by the following three equations:

$$s_{0,q} = l_{00}^{-\varsigma}\left\|\mathbf{f}_q^H\mathbf{H}_{00}\mathbf{V}_0\right\|^2,$$

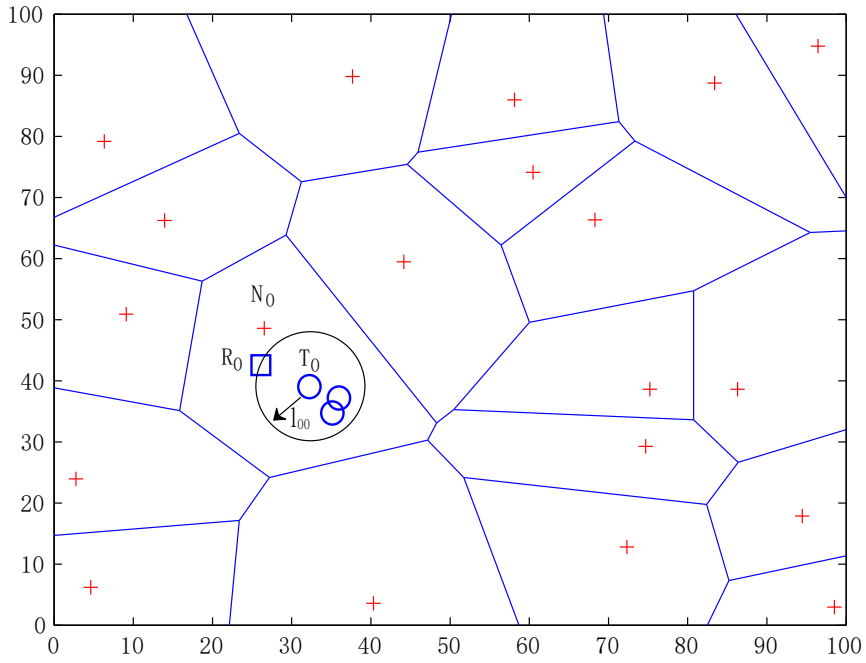


Figure 6.2: Voronoi cells of transmitters in tier-A with Poisson intensity $\lambda_A = 0.002$.

$$I_c = \left\| \mathbf{f}_q^H \sum_{L_i \in \Phi_A} l_{0L_i}^{-\varsigma/2} \mathbf{H}_{0L_i} \mathbf{V}_{L_i} \mathbf{s}_{L_i} \right\|^2 ,$$

$$I_{inter} = \left\| \mathbf{f}_q^H \sum_{T_i \in \Phi_B / \phi_0} l_{0i}^{-\varsigma/2} \mathbf{H}_{0i} \mathbf{V}_i \mathbf{s}_i \right\|^2 ,$$

where $\mathbf{f}_q$ is the q-th column of $\mathbf{U}_0$. More specific realisation of $\mathbf{f}_q$ refers to Section 2.2.3.

The total interference at the q-th data stream of R_0, denoted as I_q, is sum of the two different interferences: the inter-cluster interference I_{inter} and the cross-tier interference I_c. By further splitting the cross-tier interference into two sub-interferences: the nearest cross-tier interference I_n and the remaining cross-tier interference I_A, the total interference at the q-th data stream of R_0 is given by,

$$I_q = I_{inter} + \underbrace{I_n + I_A}_{=I_c} .$$

If the channel entries are i.i.d. $\mathcal{CN}(0,1)$ and R_0 performs single-stream decoding as in [9], the signal power is exponentially distributed with unit mean, i.e., $s_{0,q} \sim \mathrm{Exp}(1)$. Therefore, the outage probability, which is the probability that the SIR is larger than the pre-specified threshold

θ, can be expressed as,

$$\epsilon = \Pr\left(\frac{s_{0,q}}{I_q} < \theta\right) = 1 - \mathbb{E}_{I_q}\left[\exp(-\theta l_{00}^{\varsigma} I_q)\right].$$

Note that as L_0 is the nearest tier-A transmitter to R_0, l_{iL_0} is correlated with l_{iL_i}, where $L_i \neq L_0$, which leads to a correlation between the nearest cross-tier interference and the renaming cross-tier interference. However, as the remaining cross-tier interference is marginalised in this chapter, I_n and I_A are mutually independent. With properties of exponential function, the throughput, denoted as $\mathcal{T}$, can be expressed as,

$$\begin{aligned}\mathcal{T} =& KR\bar{d}_B \underbrace{\mathbb{E}\left[\exp(-\theta l_{00}^{\varsigma} I_q)\right]}_{p_s},\\ =& KR\bar{d}_B \underbrace{\mathbb{E}\left[\exp(-\theta I_n)\right]}_{p_s^n} \underbrace{\mathbb{E}\left[\exp(-\theta I_A)\right]}_{p_s^A} \underbrace{\mathbb{E}\left[\exp(-\theta I_{inter})\right]}_{p_s^{inter}},\end{aligned}$$

where p_s denotes the overall success probability which is equal to $1-\epsilon$. Symbols p_s^n, p_s^A and p_s^{inter} denote the success probabilities affected by the nearest cross-tier interference, the remaining cross-tier interference and the inter-cluster interference, respectively. In order to evaluate the performance of the IA approach in the proposed heterogeneous network, throughput, which is defined as the rate of successful message delivery over a communication system, as in (2.5), is selected as the metric. It can seen that throughput not only considers the success probability that is affected by interference, but also includes the multiplexing gain that can be enhanced by the IA approach. Therefore, it is an ideal metric to evaluate the tier-B performance in this chapter.

6.3 Interference Mitigation

In this section, the nearest cross-tier interference, i.e., the interference from L_0, is decomposed at first, which leads to the feasibility condition of the IA approach. Afterwards, transmitters in the cluster ϕ_0 will design the distance-dependent IA approach to maximise the throughput where the receivers will utilise the decorrelators to extract the desired signals.

6.3.1 Nearest Cross-Tier Interference Decomposition

The interference signal from the nearest cross-tier transmitter L_0 received at R_0 is given by,

$$l_{0L_0}^{-\varsigma/2}\mathbf{U}_0^H\mathbf{H}_{0L_0}\mathbf{V}_{L_0}\mathbf{s}_{L_0},$$

where $\mathbf{s}_{L_0} = [s_{L_0}^{(1)}\ s_{L_0}^{(2)}\ \dots\ s_{L_0}^{(\bar{d}_A)}]^T$ and the matrix $\mathbf{H}_{0L_0}\mathbf{V}_{L_0}$ can be rewritten as $[\mathbf{v}_1\ \mathbf{v}_2\ \dots\ \mathbf{v}_{\bar{d}_A}]$, where $\mathbf{v}_i \in \mathbb{C}^{N_B\times 1}$. Intuitively, $\mathbf{v}_i$ is the beamforming vector for the i-th symbol in $\mathbf{s}_{L_0}$. Recall that each column of $\mathbf{U}_0$ is the decorrelator for each data stream, i.e., $\mathbf{U}_0 = [\mathbf{f}_1\ \mathbf{f}_2\ \dots\ \mathbf{f}_{\bar{d}_B}]$. Therefore, the interference signal from L_0 ignoring the path-loss effect can be written as,

$$\mathbf{U}_0^H\mathbf{H}_{0L_0}\mathbf{V}_{L_0}\mathbf{s}_{L_0} = \begin{bmatrix} \mathbf{f}_1^H\mathbf{v}_1 & \mathbf{f}_1^H\mathbf{v}_2 & \cdots & \mathbf{f}_1^H\mathbf{v}_{\bar{d}_A} \\ \mathbf{f}_2^H\mathbf{v}_1 & \mathbf{f}_2^H\mathbf{v}_2 & \cdots & \mathbf{f}_2^H\mathbf{v}_{\bar{d}_A} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{f}_{\bar{d}_B}^H\mathbf{v}_1 & \mathbf{f}_{\bar{d}_B}^H\mathbf{v}_2 & \cdots & \mathbf{f}_{\bar{d}_B}^H\mathbf{v}_{\bar{d}_A} \end{bmatrix} \begin{bmatrix} s_{L_0}^{(1)} \\ s_{L_0}^{(2)} \\ \vdots \\ s_{L_0}^{(\bar{d}_A)} \end{bmatrix}.$$

Each entry $\mathbf{f}_i^H\mathbf{v}_j$ is chi-squared distributed with *two* DoF for two reasons. The vectors $\mathbf{v}_i, i = 1, 2, \dots, \bar{d}_A$ include the complex Gaussian distributed channel entries, the squared sum of the real and imaginary parts of which lead to chi-squared distribution. Moreover, because all the columns of $\mathbf{U}_0$ and $\mathbf{V}_{L_0}$ have unit norm, they have no impact on the signal distribution. Therefore, multiplying the effective channel $\mathbf{H}_{0L_0}\mathbf{V}_{L_0}$ with $\mathbf{f}_q^H$, the interference received at the q-th data stream of R_0 is obtained, the power of which is chi-squared distributed with $2\bar{d}_A$ DoF. If the interference from the j-th data stream from L_0 is eliminated at the i-th data stream of R_0, then $\mathbf{f}_i^H\mathbf{v}_j = 0, \forall i = 1, 2, \dots, \bar{d}_B$. Because all of the channel entries and symbols are i.i.d., it is possible to partially eliminate interference from L_0 at R_0. If $\bar{d}_a$ out of $\bar{d}_A$ data streams are treated as interference, the partially interference estimation can be expressed as follows:

$$\mathbf{U}_0^H\hat{\mathbf{H}}_{iL_0}\mathbf{V}_{L_0}\mathbf{\Sigma}_{\bar{d}_a} = \mathbf{0}, \tag{6.2}$$

where $\mathbf{\Sigma}_{\bar{d}_a}$ contains arbitrary $\bar{d}_a$ columns of a $\bar{d}_A \times \bar{d}_A$ identity matrix, denoted as $\mathbf{I}^{(\bar{d}_A)}$.

Intuitively, when the distance from R_0 to L_0 is large, the interference power from L_0 overheard by R_0 can be treated as shot noise without sacrificing much SIR at R_0. However, as the distance decreases, the interference cannot be regarded as noise and it is reasonable to partially eliminate

the interference. Hence, there exists a trade-off between treating interference from L_0 as noise and treating it as interference, i.e., a trade-off between reducing the SIR of each data stream of R_0 and reducing the number of data streams from T_0. Because the trade-off is related to the distance, the optimal solution to strike the balance should consider the distance information, i.e., IA approach should be implemented in a distance-dependent way.

6.3.2 Feasibility Condition of Interference Alignment

Signal Space at R_0

Because of the independence between all of the channel entries, if IA beamforming is not considered, $\bar{d}_a$ data streams from L_0 will occupy a $\bar{d}_a$-dimensional subspace at R_0 and each intra-cluster interference will occupy a $\bar{d}_B$-dimensional subspace. As discussed in Section 2.2.5, when IA is implemented, the interference is overlapped as much as possible. In the extreme case, i.e., when the number of antennas is enough to perfectly align all the interference, the signal space at R_0 can be demonstrated by Figure 6.3. In this figure, each rectangle represents one SRDoF at R_0. The blue rectangles represent the SRDoF exploited to decode the desired private symbols from T_0. The green ones and the red ones are the SRDoF utilised to cancel the intra-cluster interference and the nearest cross-tier interference, respectively, where the intra-cluster interference is perfectly aligned with part of the nearest cross-tier interference.

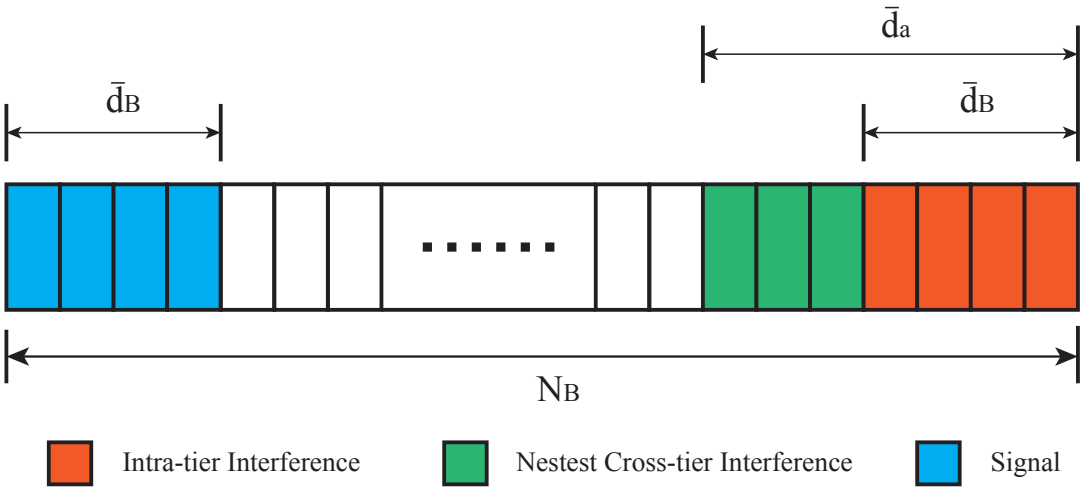


Figure 6.3: Signal space at R_0 when interference are maximum overlapped.

However, in most practical cases, the number of antennas is not enough to perfectly align all the

interference even though the feasibility condition is satisfied. For example, in a 5-user IC where each transmitter and receiver is equipped with *six* antennas. According to the feasibility condition (2.4), only *two* data streams can be transmitted. That is to say, the desired signal occupies a 2-dimensional subspace and the interference signal occupies a 3-dimensional subspace (if perfectly aligned, the interference should occupy a 2-dimensional subspace). Therefore, it is assumed that the desired signal at R_0 occupies a $\bar{d}_B$-dimensional subspace and the interference signal occupies the remaining $(N_B - \bar{d}_B)$-dimensional subspace.

Feasibility Condition

The feasibility condition is the main factor that limits the effectiveness of MIMO IA. As proven in [35], according to Bezout's theorem, IA is feasible only when the number of variables is no less than the number of equations that are required by the IA beamforming scheme.

To obtain the number of variables in any beamforming matrix $\mathbf{V}_i$, the matrices should be rewritten into the following form to guarantee the linear independence between all of the columns, i.e.,

$$\mathbf{V}_i = \begin{bmatrix} \mathbf{I}^{(\bar{d}_B)} \\ \\ \tilde{\mathbf{V}}_i \end{bmatrix},$$

where it is straightforward that $\tilde{\mathbf{V}}_i$ is a $\bar{d}_B \times (N_B - \bar{d}_B)$ matrix and there is no constraint on the design of each column. The reason to guarantee the linear independence among the columns is that two linear-dependent beamforming vectors can not provide two linear independent observations of the desired symbols. Therefore, the effective number of variables that are provided by $\mathbf{V}_i$ is $\bar{d}_B(N_B - \bar{d}_B)$. Similarly, the corresponding receiving matrices can provide the same number of variables. In total, there are $2K$ beamforming matrices and receiving matrices, excluding the beamforming matrix from L_0 because it will not be redesigned to cancel its interference at tier-B receivers. Therefore, the maximum number of variables provided to conduct the IA approach should be $2K\bar{d}_B(N_B - \bar{d}_B)$. The number of equations is determined by the definition of the IA approach. For example, at R_0, in order to eliminate the intra-cluster interference using the IA approach, the following IA equations must hold:

$$\mathbf{U}_0\mathbf{H}_{0i}\mathbf{V}_i = \mathbf{0}, \quad i \neq 0,\ T_i \in \phi_0, \tag{6.3}$$

$$\text{rank}\,(\mathbf{U}_0\mathbf{H}_{00}\mathbf{V}_0) = \bar{d}_B. \tag{6.4}$$

Meanwhile, it is assumed that $\bar{d}_a$ out of $\bar{d}_A$ data streams from L_0 are eliminated at R_0, where $\bar{d}_a \leq \min\{N_B, \bar{d}_A\}$. As discussed before, the additional equation that must be satisfied is given by (6.2).

As all of the channel entries are i.i.d., the rank equation (6.4) should be satisfied automatically. The effective number of equations included in (6.3) and (6.2) for each receiver is $(K-1)\bar{d}_B^2 + \bar{d}_a\bar{d}_B$. Adding up all the equations for the K receivers corresponding to the K transmitters in ϕ_0, the total number of equations equals $K(K-1)\bar{d}_B^2 + K\bar{d}_a\bar{d}_B$. It is well-known that IA is feasible only when the number of equations is no larger than the number of variables, which, in fact means the following equation:

$$2Kd_B(N_B - \bar{d}_B) \geq K(K-1)\bar{d}_B^2 + K\bar{d}_a\bar{d}_B \quad \Longrightarrow \quad \bar{d}_B \leq \frac{2N_B - \bar{d}_a}{K+1}.$$

This inequality presents an upper bound on the data streams transmitted by tier-B transmitters so that the IA approach can be feasibly implemented. In order to exploit all of the possible DoF available, $\bar{d}_B$ should be selected to satisfy the equality, i.e, $\bar{d}_B = \frac{2N_B - \bar{d}_a}{K+1}$. Otherwise there is still unexploited multiplexing gain.

Another constraint of $\bar{d}_a$ and $\bar{d}_B$ is that the sum of the dimensions occupied by the interference and the desired signal can not exceed the whole dimension of signal space N_B. Otherwise, the interference signal and the desired signal will be partially overlapping with each other. Thus,

$$\bar{d}_B + \max\{\bar{d}_B, \bar{d}_a\} \leq N_B \Longrightarrow \max\{4N_B - 2\bar{d}_a, 2N_B + K\bar{d}_a\} \leq N_B(K+1).$$

Finally, the constraint of $\bar{d}_a$ can be equivalently written as,

$$\bar{d}_a \leq \min\left\{\frac{N_B(K-1)}{K}, \bar{d}_A\right\}. \tag{6.5}$$

This inequality shows the upper bound of the number of data streams from L_0 that can be feasibly cancelled at R_0.

6.3.3 Distance-Dependent Interference Alignment

As shown in the previous subsection, $\bar{d}_a$ out of $\bar{d}_A$ data streams from L_0 are placed in the interference subspace (treated as interference) at R_0 using the IA approach and the remaining $\bar{d}_A - \bar{d}_a$ data streams are treated as shot noise. The optimal value of $\bar{d}_a$ is investigated so that the

throughput can be maximised. With the receiving matrix, the IA equations that have interference from ϕ_0 fully eliminated and interference from L_0 partially eliminated are as follows:

$$\begin{aligned}\mathbf{U}_0^H \hat{\mathbf{H}}_{0L_0} \mathbf{V}_{L_0} \mathbf{\Sigma}_{\bar{d}_a} &= \mathbf{0}, \\ \mathbf{U}_0 \mathbf{H}_{0i} \mathbf{V}_i &= \mathbf{0}, \quad \forall T_i \in \phi_0, i \neq 0 \\ \text{rank}(\mathbf{U}_0 \mathbf{H}_{00} \mathbf{V}_0) &= \bar{d}_B.\end{aligned}$$

Because of the imperfect CSI (with the variance of estimation error being ε^2), each of the $\bar{d}_a$ data streams that are treated as interference still causes interference with power reduced by ε^2 compared to the interference caused by each of the other $\bar{d}_A - \bar{d}_a$ data streams. The interference power at the q-th data stream of R_0 is, therefore, represented as,

$$\begin{aligned}I_n &= l_{0L_0}^{-\varsigma} \|\mathbf{f}_q^H \mathbf{H}_{0L_0} \mathbf{V}_{L_0} \mathbf{s}_{L_0}\|^2 \\ &= P_A l_{0L_0}^{-\varsigma} \varepsilon^2 \|\mathbf{f}_q^H \tilde{\mathbf{H}}_{0L_0} \mathbf{V}_{L_0} \mathbf{\Sigma}_{\bar{d}_a}\|^2 + P_A l_{0L_0}^{-\varsigma} \|\mathbf{f}_q^H \mathbf{H}_{0L_0} \mathbf{V}_{L_0} \tilde{\mathbf{\Sigma}}_{\bar{d}_a}\|^2 \\ &= I_n^{(1)} + I_n^{(2)},\end{aligned}$$

where $\tilde{\mathbf{\Sigma}}_{\bar{d}_a}$ is the complementary $\bar{d}_A - \bar{d}_a$ columns of $\mathbf{\Sigma}_{\bar{d}_a}$ in the identity matrix. As the channel is complex Gaussian distributed, $I_n^{(1)}$ and $I_n^{(2)}$ are chi-squared distributed with $2\bar{d}_a$ and $2(\bar{d}_A - \bar{d}_a)$ DoF respectively.

Theorem 9 *The success probability at each data stream of receiver R_0 affected by the nearest cross-tier interferer L_0 under Rayleigh fading is given by,*

$$p_s^n = \left(P_A l_{0L_0}^{-\varsigma} \theta l_{00}^{\varsigma} + 1\right)^{-d},$$

where L_0 is assumed to transmit d data streams.

Proof: See Appendix D.

The interference from L_0 in this chapter actually consists of two additive interference sources: the $\bar{d}_a$ data streams that are treated as interference and remaining data streams that are treated as shot noise. Hence, the success probability affected by interference from L_0 is the product of the success probabilities affected by the two interference sources. Based on Theorem 9, the success probability affected by L_0 can be considered as a product of the two success probabilities, one is affected by the interference from $\bar{d}_a$ data streams with power $\varepsilon^2 P_A$ and the other is affected by the interference from $\bar{d}_A - \bar{d}_a$ data streams with symbol power P_A. Thus, the success probability of

each data stream of R_0 affected by L_0 is given by,

$$p_s^n = \exp\left(-\theta l_{00}^{\varsigma} I_n^{(1)}\right) \exp\left(-\theta l_{00}^{\varsigma} I_n^{(2)}\right) = \left(P_A l_{0L_0}^{-\varsigma} \varepsilon^2 \theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_a} \left(P_A l_{0L_0}^{-\varsigma} \theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_A + \bar{d}_a}.$$

Because $\bar{d}_B = \frac{2N_B - \bar{d}_a}{K+1}$, the corresponding throughput can now be obtained as,

$$\mathcal{T} = KRp_s^A p_s^{inter} \left(\frac{2N_B - \bar{d}_a}{K+1}\right) \left(P_A l_{0L_0}^{-\varsigma} \varepsilon^2 \theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_a} \left(P_A l_{0L_0}^{-\varsigma} \theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_A + \bar{d}_a}. \quad (6.6)$$

From this equation it can be seen that the success probability is affected by the distance from L_0 to R_0 which in practice, R_0 may or may not know. Therefore, $\bar{d}_a$ that maximises the throughput is found in two scenarios: l_{0L_0} is only known statistically and l_{0L_0} is accurately known. Generally speaking, a large $\bar{d}_a$ leaves a lower dimensional subspace for the desired signals because $\bar{d}_B = \frac{2N_B - \bar{d}_a}{K+1}$, but the SIR of each data stream increases. For small $\bar{d}_a$, there will be a higher dimensional subspace for the desired signals, but the SIR at each data stream will be lower. It is worth mentioning that the expected throughput is maximised independently with given p_s^A and p_s^{inter}. The reason behind this is that transmitter-receiver pairs in a cluster is maximising the throughput for the cluster wherein p_s^A and p_s^{inter} are not controllable. The optimal number of data streams from L_0 that should be treated as interference is denoted as $\hat{d}_a$.

R_0 Has Statistical Knowledge of l_{0L_0}

This condition implies that $\hat{d}_a$ is chosen according to statistical information. For example, based on the property of the PPP with intensity λ_A, according to [94], the PDF of the distance from R_0 to L_0 is given by,

$$f_{L_0}(x) = 2\lambda_A \pi x e^{-\pi \lambda_A x^2}.$$

Considering $\bar{d}_B = \frac{2N_B - \bar{d}_a}{K+1}$, the expected value of the throughput is given by,

$$\begin{aligned}\mathbb{E}[\mathcal{T}] &= KRp_s^A p_s^{inter} \bar{d}_B \int_0^\infty \left(P_A x^{-\varsigma} \varepsilon^2 \theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_a} \left(P_A x^{-\varsigma} \theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_A + \bar{d}_a} f_{L_0}(x) \mathrm{d}x \\ &= 2\lambda_A \pi KRp_s^A p_s^{inter} \left(\frac{2N_B - \bar{d}_a}{K+1}\right) \int_0^\infty \left(P_A x^{-\varsigma} \varepsilon^2 \theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_a} \times \\ &\qquad \left(P_A x^{-\varsigma} \theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_A + \bar{d}_a} x e^{-\lambda_A \pi x^2} \mathrm{d}x.\end{aligned} \quad (6.7)$$

Defining all the terms containing $\bar{d}_a$ as $\mathcal{Q}_8$, the following equation can be obtained:

$$\mathcal{Q}_8 = (2N_B - \bar{d}_a)\left(P_A x^{-\varsigma}\varepsilon^2\theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_a}\left(P_A x^{-\varsigma}\theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_A+\bar{d}_a}.$$

It can be easily proven that the second order derivative of $\mathcal{Q}_8$ is negative. Thus, $\mathcal{Q}_8$ is a concave function with respect to $\bar{d}_a$, which implies that $\mathbb{E}[\mathcal{T}]$ is concave because integral operation maintains concavity. As a result, the value of $\bar{d}_a$ that maximises the throughput, denoted as $\hat{d}_a$, is able to be found by the following expressions,

$$\begin{aligned} \arg\max_{\bar{d}_a} \quad & \mathbb{E}[\mathcal{T}], \\ \text{s.t.} \quad & \bar{d}_a \leq \min\left\{\frac{N_B(K-1)}{K}, \bar{d}_A\right\}. \end{aligned}$$

After obtaining $\hat{d}_a$, the closed-form expression of (6.7) is not available for general ς. However, it is available when the quality of the channel from R_0 to L_0 is high, i.e., $\varepsilon = 0$ and fixed ς. For example, when $\varepsilon = 0$ and $\varsigma = 4$, the integral part of (6.7) is given by,

$$\int_0^{\infty}\left(P_A l^{-4}\theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_A+\hat{d}_a} l e^{-\lambda_A \pi l^2} dl = \frac{\sqrt{\theta l_{00}^{\varsigma} P_A}\,\mathcal{G}_{1,3}^{3,1}\left(\frac{(\pi\lambda_A)^2\theta l_{00}^{\varsigma} P_A}{4}\middle|\begin{matrix}\hat{d}_a-\bar{d}_A+\frac{1}{2} \\ -\frac{1}{2},0,\frac{1}{2}\end{matrix}\right)}{4\sqrt{\pi}\Gamma(\bar{d}_A - \bar{d}_a)},$$

where $\mathcal{G}(\cdot)$ is the Meijer-G function. The closed-form expression of an upper bound of (6.7) is available. Specifically, as $f_{L_0}(x)$ is a PDF, the integral part can be thought of as $\mathbb{E}_x\left[\left(P_A x^{-\varsigma}\varepsilon^2\theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_a}\left(P_A x^{-\varsigma}\theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_A+\bar{d}_a}\right]$, which is a concave function. With Jensen's inequality, the upper bound of the throughput is given by,

$$\mathbb{E}[\mathcal{T}] \leq K R p_s^A p_s^{inter}\left(\frac{2N_B - \hat{d}_a}{K+1}\right)\left(2^{\varsigma} P_A \lambda_A^{\varsigma/2}\varepsilon^2\theta l_{00}^{\varsigma} + 1\right)^{-\hat{d}_a}\left(2^{\varsigma} P_A \lambda_A^{\varsigma/2}\theta l_{00}^{\varsigma} + 1\right)^{-\bar{d}_A+\hat{d}_a}$$

R_0 Has Accurate Knowledge of l_{0L_0}

In this scenario, $\hat{d}_a$ is chosen according to accurate l_{0L_0} instead of statistical knowledge. The optimal value $\hat{d}_a$ is the value of $\bar{d}_a$ that makes the first-order derivative of $\mathcal{Q}_8$ equal to zero. The

first-order derivative is given by,

$$\begin{aligned}\mathcal{Q}_8' = &- \left(P_A x^{-\varsigma}\theta l_{00}^{\varsigma}+1\right)^{\bar{d}_a-\bar{d}_A}\left(P_A x^{-\varsigma}\varepsilon^2\theta l_{00}^{\varsigma}+1\right)^{-\bar{d}_a} \\ &+ \left(P_A x^{-\varsigma}\theta l_{00}^{\varsigma}+1\right)^{\bar{d}_a-\bar{d}_A}\left(P_A x^{-\varsigma}\varepsilon^2\theta l_{00}^{\varsigma}+1\right)^{-\bar{d}_a}\ln\left(P_A x^{-\varsigma}\theta l_{00}^{\varsigma}+1\right)(-\bar{d}_a+2N_B) \\ &- \left(P_A x^{-\varsigma}\theta l_{00}^{\varsigma}+1\right)^{\bar{d}_a-\bar{d}_A}\left(P_A x^{-\varsigma}\varepsilon^2\theta l_{00}^{\varsigma}+1\right)^{-\bar{d}_a}\ln\left(P_A x^{-\varsigma}\varepsilon^2\theta l_{00}^{\varsigma}+1\right)(-\bar{d}_a+2N_B).\end{aligned}$$

By setting $\mathcal{Q}_8' = 0$, the following equation can be obtained:

$$\hat{d}_a = \arg\max_{\bar{d}_a}\mathbb{E}[\mathcal{T}] = 2N_B - \left[\ln\left(\frac{\left(P_A l_{0L_0}^{-\varsigma}\theta l_{00}^{\varsigma}+1\right)}{\left(P_A l_{0L_0}^{-\varsigma}\varepsilon^2\theta l_{00}^{\varsigma}+1\right)}\right)\right]^{-1}, \tag{6.8}$$

where $\hat{d}_a \leq \min\left\{\frac{N_B(K-1)}{K}, \bar{d}_A\right\}$.

Remark: In the case when $\varepsilon \to 1$, the logarithmic function tends to be *zero* and so is $\hat{d}_a$. This is because when the quality of CSI is very bad, there is only slight difference between considering the interference signal as interference or as shot noise. Hence, there is no need to expend extra variables to align the interference to L_0. That is to say, this expression has already taken the feasibility condition of IA into consideration.

As $\mathcal{Q}_8$ is a concave function, using $\hat{d}_a$ as the number of data streams from L_0 that are treated as interference leads to the maximum $\mathcal{Q}_8$. In this way, the analytic expression of the maximised throughput can be written as,

$$\begin{aligned}\mathbb{E}[\mathcal{T}] = 2\lambda_A\pi k R p_s^A p_s^{inter}\int_0^{\infty}&\left(\frac{2N_B-\hat{d}_a}{K+1}\right) \\ &\left(P_A x^{-\varsigma}\varepsilon^2\theta l_{00}^{\varsigma}+1\right)^{-\hat{d}_a}\left(P_A x^{-\varsigma}\theta l_{00}^{\varsigma}+1\right)^{-\bar{d}_A+\hat{d}_a} x e^{-\lambda_A\pi l^2}\mathrm{d}x.\end{aligned}$$

6.4 Outage Probability and Throughput

In this section, the success probabilities that are affected by the remaining interference from tier-A transmitters as well as the inter-cluster interference from tier-B transmitters are derived using stochastic geometry technique. Recall that the two success probabilities are denoted as p_s^A and p_s^{inter}, respectively.

6.4.1 Remaining Cross-Tier Interference

The remaining cross-tier interference is the interference caused by the tier-A transmitters except for L_0, which is essentially a PPP without the nearest point. The outage probability at the q-th data stream of R_0 affected by the remaining cross-tier interference is given by,

$$1-p_s^A=\Pr\left(\frac{l_{00}^{-\varsigma}\|\mathbf{f}_q^H\mathbf{H}_{00}\mathbf{V}_i\|^2}{\sum_{\substack{L_i\in\Phi_A,\\ L_i\neq L_0}}P_A l_{0L_i}^{-\varsigma}\|\mathbf{f}_q^H\mathbf{H}_{0L_i}\mathbf{V}_{L_i}\|^2}<\theta\right).$$

For Rayleigh fading, the corresponding success probability is given by,

$$p_s^A=\mathcal{L}_{I_A}(s)|_{s=\theta}=\mathbb{E}\left[\exp\left(-P_A\theta l_{00}^{\varsigma}\sum_{\substack{L_i\in\Phi_A,\\ L_i\neq L_0}}l_{0L_i}^{-\varsigma}\|\mathbf{f}_q^H\mathbf{H}_{0L_i}\mathbf{V}_{L_i}\|^2\right)\right].$$

For any measure $\nu:\mathcal{A}_l\rightarrow[0,\infty)$, where $\mathcal{A}_l$ is the area with distance to R_0 larger than l_{0L_0}, the PGF of PPP, denoted as $\mathfrak{G}_{PPP}$, is given by,

$$\mathfrak{G}_{PPP}[v(x)]=\exp\left(-\int_{\mathcal{A}_l}(1-v(x))\Lambda(\mathrm{d}x)\right), \tag{6.9}$$

where $\Lambda(\cdot)$ is the intensity measure introduced in Section 2.3.3. For tier-A transmitters with intensity λ_A and $v(l)=\exp\left(-\frac{\theta l_{00}^{\varsigma}h\bar{P}_A l^{-\varsigma}}{2}\right)$, the success probability affected by all tier-A transmitters except L_0 can be expressed as,

$$p_s^A=\exp\left(-2\pi\lambda_A\mathbb{E}_h\left[\int_0^{\infty}\int_x^{\infty}\left(1-\exp\left(-\frac{\theta l_{00}^{\varsigma}h\bar{P}_A l^{-\varsigma}}{2}\right)\right)x f_{L_0}(l)\mathrm{d}l\mathrm{d}x\right]\right). \tag{6.10}$$

Remark: Symbol h denotes the effective fading coefficient and is chi-squared distributed with $2\bar{d}_A$ DoF. In tier-A, all of the transmitters are independently located which means that all of the transmitters except L_0 itself are evenly distributed outside L_0. This property of PPP can be made use of to marginalise the interference from $L_i,\forall i\neq 0$. Specifically, as seen in the inner integral, the integration of the interference is from x, which is the integral variable for the distance from R_0 to L_0. By further considering the PDF of the distribution of L_0, an accurate expression of the marginalised interference power from the remaining cross-tier interferers is obtained. Therefore, the success probability affected by the cross-tier interference can be divided in to two independent components: nearest cross-tier interference and marginalised remaining cross-tier interference.

Theorem 10 *With all the transmitters in tier-A being PPP, the success probability affected by*

tier-A interferers excluding the nearest one at each data stream of R_0 is given by,

$$p_s^A = \exp\left(1 + \frac{2\pi\lambda_A (P_A\theta)^{\frac{2}{\varsigma}} l_{00}^2 \Gamma(\bar{d}_A + \frac{2}{\varsigma})\Gamma(-\frac{2}{\varsigma})}{\varsigma\Gamma(\bar{d}_A)} - \pi\lambda_A \int_0^\infty e^{-\pi\lambda_A t}(1 + P_A\theta l_{00}^{\varsigma} t^{-\varsigma/2})^{-\bar{d}_A} dt\right).$$

Proof: See Appendix E

Although Theorem 10 is valid for any $\varsigma > 2$, the integral part can not be simplified directly. However, for any particular value of ς, closed-form expression is available. For example, when $\varsigma = 4$,

$$\int_0^\infty e^{-\pi\lambda_A t}(1 + P_A\theta l_{00}^{\varsigma} t^{-2})^{-\bar{d}_A} \mathrm{d}t = \frac{\sqrt{\theta l_{00}^{\varsigma} P_A}\mathcal{G}_{1,3}^{3,1}\left(\frac{(\pi\lambda_A)^2 P_A\theta l_{00}^{\varsigma}}{4}\middle|\begin{matrix}\frac{1}{2}-\bar{d}_A \\ \\ -\frac{1}{2},0,\frac{1}{2}\end{matrix}\right)}{2\sqrt{\pi}\Gamma(\bar{d}_A)}.$$

Although closed-form expression is not available with general ς, the upper bound can be found by regarding $\pi\lambda_A \exp(-\pi\lambda_A t)$ as the PDF of t. Therefore, the integral part in the exponential function of Theorem 10 can be interpreted as $\mathbb{E}[(1 + P_A\theta l_{00}^{\varsigma} t^{-\varsigma/2})^{-\bar{d}_A}]$, which is concave in terms of t. Using Jensen's inequality, then,

$$\mathbb{E}[(1 + P_A\theta l_{00}^{\varsigma} t^{-\varsigma/2})^{-\bar{d}_A}] \geq \left(1 + P_A\theta l_{00}^{\varsigma}\mathbb{E}[t]^{-\varsigma/2}\right)^{-d_A} = \left(1 + P_A\theta l_{00}^{\varsigma}(\pi\lambda_A)^{\varsigma/2}\right)^{-d_A}.$$

Therefore, an upper bound of the success probability affected by the remaining cross-tier interference can be written as,

$$p_s^A \leq \exp\left(1 + \frac{2\pi\lambda_A (P_A\theta)^{\frac{2}{\varsigma}} l_{00}^2 \Gamma(\bar{d}_A + \frac{2}{\varsigma})\Gamma(-\frac{2}{\varsigma})}{\varsigma\Gamma(\bar{d}_A)} - \left(1 + P_A\theta l_{00}^{\varsigma}(\pi\lambda_A)^{\varsigma/2}\right)^{-\bar{d}_A}\right).$$

6.4.2 Inter-Cluster Interference

In tier-B, transmitters are distributed as clustered process with parent points intensity λ_B and each cluster includes K transmitters uniformly distributed within range r_a of the parent point. The success probability caused by the inter-cluster interference at R_0 is represented by,

$$p_s^{inter} = \Pr\left(\frac{l_{00}^{-\varsigma}\|\mathbf{f}_q^H\mathbf{H}_{00}\mathbf{V}_i\|^2}{\sum_{T_i\in\Phi_A/\phi_0} l_{0i}^{-\varsigma}\|\mathbf{f}_q^H\mathbf{H}_{0i}\mathbf{V}_i\|^2} \geq \theta\right).$$

The success probability with Poisson clustered distributed interferers can be expressed as product of intra-cluster interference and inter-cluster interference [85]. In this chapter, the intra-cluster interference is thought to be completely eliminated by the decorrelator with perfect intra-cluster CSI. Therefore, it is only necessary to derive the success probability affected by inter-cluster interference.

Considering the specific Poisson clustered process of the tier-B transmitters in this chapter, the PGF should be given by,

$$\mathfrak{G}_{PCP}[\nu(x)] = \exp\left(-\lambda_B \int_{\mathbb{R}^2} 1 - \left(\int_{b(o,r_a)} v(x+y) f_B(y) \mathrm{d}y\right)^K \mathrm{d}x\right),$$

where the same measure $\nu(x)$ is selected as that in (6.9) towards the success probability and o denotes the origin. Therefore, the success probability affected by the inter-cluster transmitters is given by,

$$\begin{aligned} p_s^{inter} &= \exp\left(-\lambda_B \int_{\mathbb{R}^2} 1 - \left(\frac{1}{\pi r_a^2} \int_{b(o,r_a)} \mathbb{E}_h\left[\exp\left(-\frac{\theta l_{00}^{\varsigma} h \|x+y\|^{-\varsigma}}{2}\right)\right] \mathrm{d}y\right)^K \mathrm{d}x\right) \\ &= \exp\left(-\lambda_B \int_{\mathbb{R}^2} 1 - \left(\frac{1}{\pi r_a^2} \int_{b(o,r_a)} \int_0^{\infty} \exp\left(-\frac{\theta l_{00}^{\varsigma} h \|x+y\|^{-\varsigma}}{2}\right) f_{\chi^2}(h, 2\bar{d}_B) \mathrm{d}h \mathrm{d}y\right)^K \mathrm{d}x\right) \\ &= \exp\left(-\lambda_B \int_{\mathbb{R}^2} 1 - \left(\frac{1}{\pi r_a^2} \int_{b(o,r_a)} \left(\frac{\|x+y\|^{\varsigma}}{\|x+y\|^{\varsigma} + \theta l_{00}^{\varsigma}}\right)^{\bar{d}_B} \mathrm{d}y\right)^K \mathrm{d}x\right), \end{aligned}$$

where $\frac{1}{\pi r_a^2}$ is the PDF of the uniformly distributed transmitters in each cluster, which is a ball with radius r_a. The success probability can be interpreted as,

$$p_s^{inter} = \exp\left(-\lambda_B \int_{\mathbb{R}^2} 1 - \mathbb{E}\left[\left(\frac{\|x+y\|^{\varsigma}}{\|x+y\|^{\varsigma} + \theta l_{00}^{\varsigma}}\right)^{\bar{d}_B}\right]^K \mathrm{d}x\right).$$

As discussed in the system model, the coverage range of macro BS is much larger than that of femto BS, so that the accurate expression can be approximated in the case when the cluster size is small, i.e.,

$$\lim_{r_a \to 0} \mathbb{E}\left[\left(\frac{\|x+y\|^{\varsigma}}{\|x+y\|^{\varsigma} + \theta l_{00}^{\varsigma}}\right)^{\bar{d}_B}\right] = \left(\frac{\|x\|^{\varsigma}}{\|x\|^{\varsigma} + \theta l_{00}^{\varsigma}}\right)^{\bar{d}_B}$$

Therefore, the success probability affected by interference from inter-cluster transmitters can be

approximated by,

$$\begin{aligned} p_s^{inter} &\approx \exp\left(-\lambda_B \int_{\mathbb{R}^2} 1 - \left(\frac{\|x\|^{\varsigma}}{\|x\|^{\varsigma} + \theta l_{00}^{\varsigma}}\right)^{\bar{d}_B K} \mathrm{d}x\right) \\ &= \exp\left(-2\pi\lambda_B \int_0^{\infty} 1 - \left(\frac{x^{\varsigma}}{x^{\varsigma} + \theta l_{00}^{\varsigma}}\right)^{\bar{d}_B K} \mathrm{d}x\right) \\ &= \exp\left(\frac{-K\bar{d}_B \theta^{2/\varsigma} l_{00}^2 \Gamma\left(\frac{\varsigma-2}{\varsigma}\right) \Gamma\left(\frac{2}{\varsigma} + K\bar{d}_B\right)}{(\pi\lambda_B)^{-1}\Gamma(1 + K\bar{d}_B)}\right). \end{aligned} \tag{6.11}$$

6.5 Simulation Results

The simulation results of this chapter are presented and discussed in this section. Specific simulation steps for the tier-A interference are similar to those described in Section 5.6 except that only the nearest transmitter is cancelled. For the tier-B transmitters, the parent points are generated at first. According to the number of parent points, the same number of the daughter processes are generated. The daughter processes are mutually independent and each has K transmitters evenly located within a ball $b(o, r_a)$. Finally, the coordinates of transmitters in the untranslated daughter processes are added to the coordinate of the corresponding parent point and, in this way, obtain the whole clustered process. Similarly, each simulation point is averaged over at least 3000 transmitter distribution and channel realisations. The default value for all the parameters can be found in Table 6.1.

Parameter Name	Symbol	Default Value
Path-loss exponent	ς	4
SINR threshold	θ	10
Tier-A transmitter intensity	λ_A	0.01
Tier-B parent point intensity	λ_B	0.05
Tier-A transmitter antennas	N_A	8
Tier-B transmitter antennas	N_B	8
Tier-A data streams	$\bar{d}_A$	8
Number of daughter points	K	3

Table 6.1: Default values for parameters.

Figure 6.4 shows the optimal number of data streams from L_0 that should be treated as interference

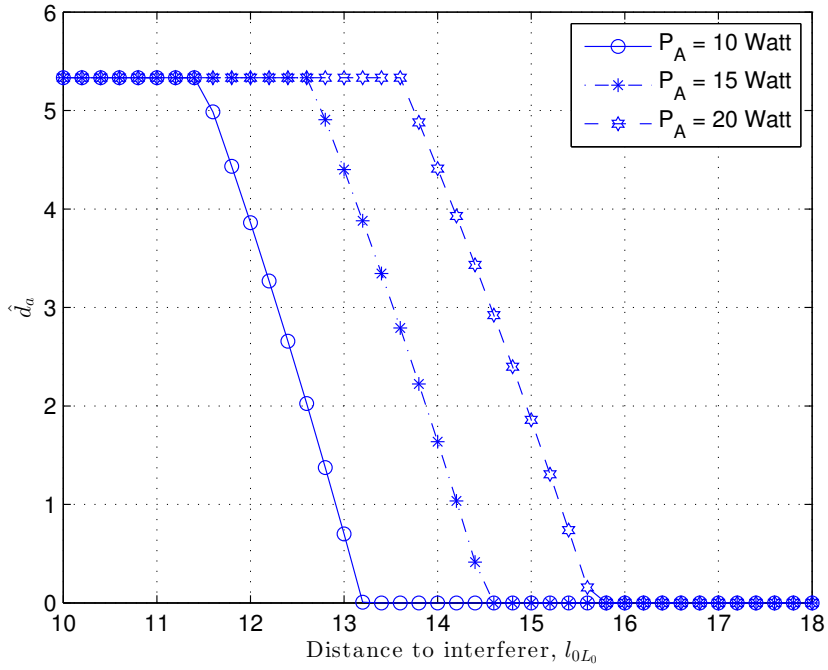


Figure 6.4: Number of data streams from L_0 that should be eliminated.

at R_0 from (6.8) versus the distance l_{0L_0}, which is assumed to be accurately known by R_0. The constraint of $\bar{d}_a$ in (6.5) is considered and $\varsigma = 3$. It can be seen that with larger P_A or smaller l_{0L_0}, interference power from each data stream of L_0 to R_0 becomes so critical that it will considerably reduce the SIR at R_0 if treated as shot noise. Therefore, more data streams from L_0 are put into the interference subspace. One important observation is that $\hat{d}_a$ will decreases to *zero* as l_{0L_0} increases. Intuitively, as there is always $\bar{d}_B$-dimensional interference subspace at each tier-B receiver, $\hat{d}_a$ should be no smaller than $\bar{d}_B$. However, Bezout's theorem states that the essence of the IA scheme is a number of equations and variables. Aligning every single additional data stream is a overhead to the beamforming or receiving matrices. Finally, when $\hat{d}_a = 0$, all the nearest cross-tier interference is treated as shot noise, which is how IA is implemented in [86–88]. Clearly, ignoring the nearest cross-tier interference is only optimal when l_{0L_0} is large.

Figure 6.5 illustrates the throughput versus transmit power P_A with $\varsigma = 3$. In order to focus on the nearest cross-tier interference, $p_s^A = p_s^{inter} = 1$ is considered. Throughput achieved with the accurate knowledge of the distance l_{0L_0} outperforms the curve with only statical knowledge. The reason behind this is that, with accurate knowledge, $\hat{d}_a$ can be selected to maximise the throughput for each transmitter distribution realisation while with statical knowledge of l_{0L_0}, $\hat{d}_a$ can only be designed to maximise the expected throughput. Interestingly, the gap between lines with accurate knowledge and statical knowledge increases as the quality of CSI increases. The reason behind this

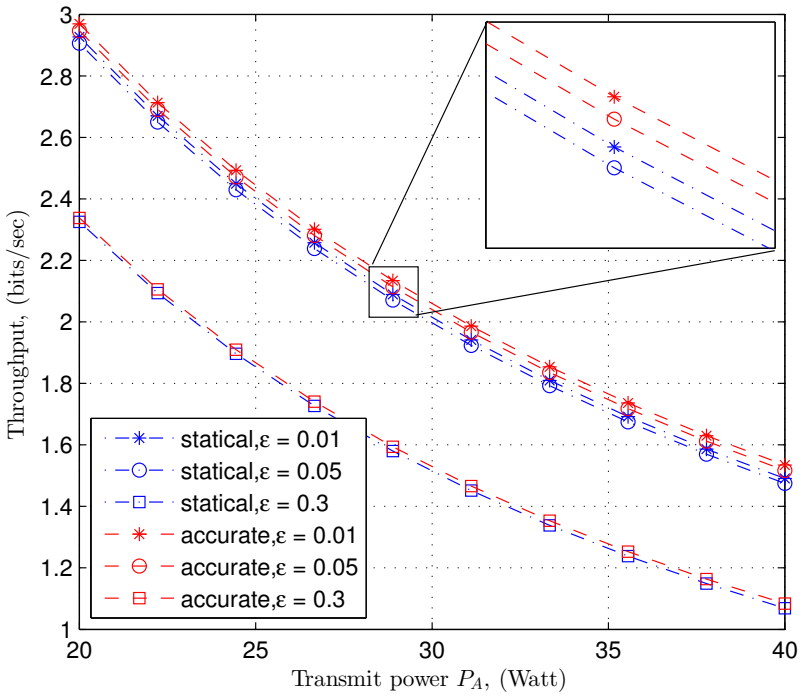


Figure 6.5: The throughput only affected by the nearest transmitter in tier-A.

is that with larger ε, convexity of (6.6) in terms of $\bar{d}_a$ increases. However, even when ε decreases to 0.01, the throughput gap with and without the accurate distance knowledge is less than 0.1 bits/s. Conclusively speaking, obtaining accurate distance knowledge to the nearest cross-tier transmitter is not significantly important. Another observation from this figure is that decreasing ε from 0.05 to 0.01 will achieve insignificant throughput gain. This observation is contrary to that in [23], which stated that even small amount of CSI error of nearby transmitters would cause severe performance loss. The reason behind this is that the proposed distance-dependent IA approach also takes the variance of CSI error into consideration, i.e., when the variance is large, fewer data streams are expected to be eliminated. In other words, the trade-off between increasing the multiplexing gain and SIR enhancement is effective.

Figure 6.6 demonstrates the success probabilities at R_0 affected only by the tier-A transmitters without the nearest one, i.e., p_s^A in Theorem 10, together with the corresponding Monte Carlo simulations. In this figure, $\varsigma = 3.5$ and $\varsigma = 4.5$ are considered, respectively. It can be seen that the success probability increases with ς, which will cause more severe path-loss through the interference signal propagation. The gap between two groups of lines increases with tier-A transmit power until a *zero* success probability is enforced. This is because the success probability is an exponential function of interference power, i.e., $\exp(-\theta I_A)$, while the interference power

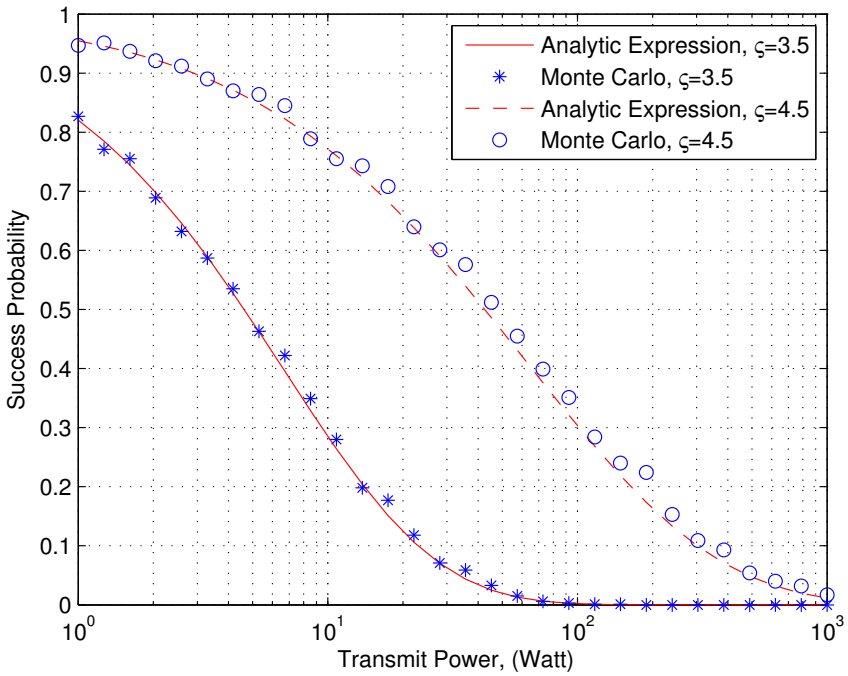


Figure 6.6: The success probability at R_0 affected by the remaining tier-A transmitters.

increases only linearly with tier-A transmit power. Moreover, this figure also shows that Theorem 10 accurately models and successfully marginalises the interference from the remaining cross-tier interferers.

The next simulation is the success probability affected only by inter-cluster interference, i.e., p_s^{inter} versus the path-loss exponent ς as shown in Figure 6.7. The condition $\lambda_B = 0.01$ and $\bar{d}_B = 2$ is considered. The success probability increases with ς because larger ς represents larger path-loss which results in smaller interference. In the analysis expression (6.11), it is assumed that all of the transmitters in a cluster are overlapped with each other, i.e., $r_a = 0$. This explains why the Monte Carlo simulation curves better approximate the line with smaller r_a. One important observation is that the success probability increases as r_a decreases. From this observation, it can be concluded that the interference power generated by a Poisson clustered network increases with the cluster size. As a result, for a clustered wireless network, shrinking the cluster size can help to reduce the interference. Finally, the gaps between the three lines decrease with smaller ς. This is because when the path-loss exponent decreases, the convexity of the path-loss function $l_{ij}^{-\varsigma}$ decreases. Hence, there will be a smaller difference between interference power from nearby and distant transmitters.

Figure 6.8 illustrates the mean throughput affected by all of the interference sources discussed

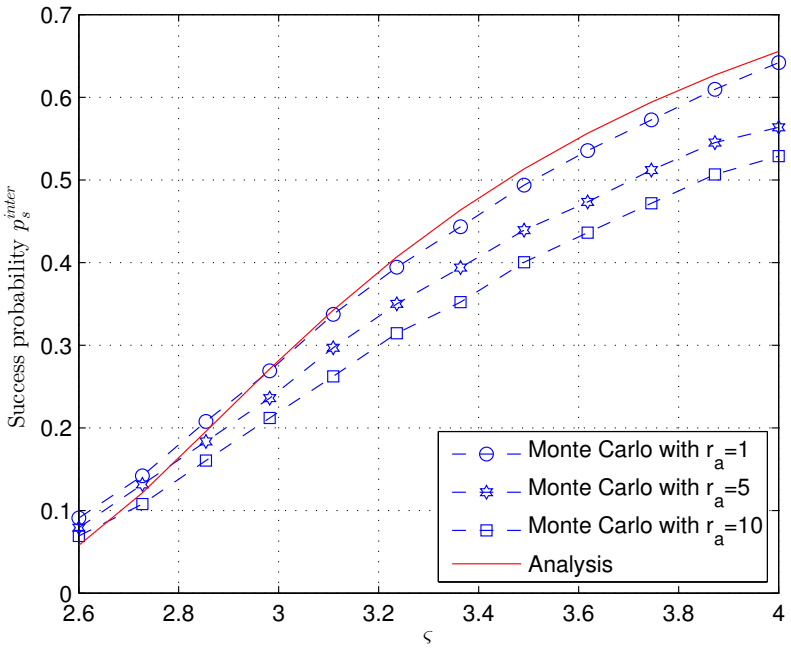
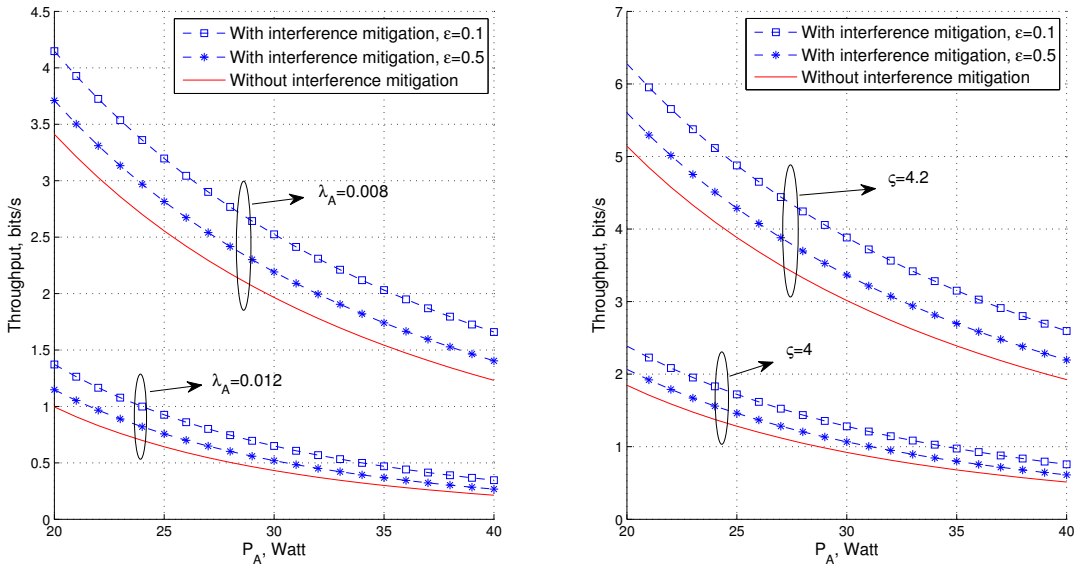


Figure 6.7: The success probability at R_0 affected by the inter-cluster transmitters.



(a) With various λ_A.

(b) With various ς.

Figure 6.8: The overall expected throughput versus tier-A transmit power.

in this chapter. Intra-cluster interference is perfectly cancelled and the interference mitigation in this figure specifically refers to mitigating the interference from the nearest cross-tier transmitter. Therefore, the lines without interference mitigation represent the IA scheme in [86–88]. The distance l_{0L_0} is assumed to be statistically known. In Figure 6.8a, $\varsigma = 4$ is fixed with varying λ_A while in Figure 6.8b, $\lambda_A = 0.01$ is fixed with varying ς. Generally, in both sub-figures, the throughput decreases with tier-A transmit power and when ϵ decreases. Specifically, in Figure 6.8a, the gain from the line without interference mitigation to the line with interference mitigation ($\epsilon = 0.1$) is 23% when $\lambda_A = 0.008$ but raises to 40% when $\lambda_A = 0.012$. This is because the expected distance from tier-A transmitters to R_0 decreases at the same rate when intensity increases, as proven in (5.2), while the interference power from one transmitter decays exponentially with its distance to R_0. Hence, when transmitter intensity increases, the interference power from the closest transmitters becomes more dominant. As the distance-dependent IA approach only deals with the nearest cross-tier interference, it has better effectiveness with larger λ_A. In Figure 6.8b, it can be observed that the throughput is highly sensitive to the path-loss exponent ς. Although increasing ς makes the nearest cross-tier interference more dominant, as in Figure 6.8a, the throughput gain decreases with larger ς. In detail, the throughput gain is 20% when $\varsigma = 4.2$ and 31% when $\varsigma = 4$. The reason behind this is that when ς increases, interference power from all of the transmitters decreases, which reduces the effectiveness of interference cancellation. Actually, when $\varsigma \to \infty$, the path-loss effect itself will eliminate all interference.

6.6 Summary

A generic two-tier heterogeneous wireless network was considered in which tier-A transmitters, with large coverage area, were Poisson distributed and tier-B transmitters, with small coverage area, were Poisson clustered distributed. Interference at tier-B receiver was categorised into three kinds: intra-cluster interference, inter-cluster interference, and cross-tier interference, where the cross-tier interference could be further divided into the nearest cross-tier interference and the remaining cross-tier interference. The conclusive remarks could be divided into two parts:

- **Interference mitigation:** A distance-dependent IA approach was proposed within the tier-B clusters in which the intra-cluster interference was aligned to a part of the interference from the nearest cross-tier transmitter. Essentially, the goal was to find the optimal part of the interference signal from the nearest cross-tier transmitter that should be treated as shot noise. The effectiveness of the distance-dependent IA approach was realised by striking a

balance between increasing the number of data streams and increasing the average SIR at each data stream. The feasibility condition of the IA approach was carefully studied using Bezout's theorem and explicitly influenced the throughput. With the feasibility condition, it was shown that, when power of the nearest cross-tier interference was low, there was no need to align intra-cluster interference to any part of it. Moreover, the comparison between the throughputs achieved with various system settings shows that: 1) Obtaining the accurate knowledge of the distance to the nearest cross-tier transmitter was not significantly important; 2) Minor channel estimation error was not harmful to the throughput, thanks to the effective trade-off between increasing multiplexing gain and the SIR enhancement; 3) The distance-dependent IA approach was more effective with a smaller path-loss exponent and a higher intensity of tier-A transmitters.

- **Stochastic interference modelling:** The remaining cross-tier interference and the inter-cluster interference were modelled with stochastic geometry technique, wherein the a correlation between the nearest cross-tier interference and remaining cross-tier interference was considered and the remaining cross-tier interference was marginalised. The near closed-form expression of the success probability affected by the remaining cross-tier interference was derived together with a closed-form upper bound. Moreover, the accurate expression of the success probability affected by the inter-cluster interference was derived as well as its closed-form approximation when the cluster size was small. Note that all of the derived expressions were verified by Monte Carlo simulations. It was showed that the influence of the path-loss exponent to the throughput is more significant with larger tier-A transmit power and shrinking the cluster size could help to reduce the mean interference power generated by Poisson clustered distributed transmitters.

Chapter 7

Conclusions and Future Work

In this thesis, advanced interference management techniques considering practical challenges in modern wireless networks were studied and evaluated. In the first chapter of this thesis, the achievable sum rate by the MAT scheme was enhanced by the probabilistic-constrained optimisation method in the finite SNR case. In the second chapter, the MAT scheme was extended to the time-correlated MIMO channels with various settings. Next, stochastic geometry technique helped to evaluate the performance of the ZF receivers in Poisson distributed wireless networks. Finally, a distance-dependent IA approach was designed for a generic 2-tier heterogeneous wireless network.

In the reminder of this chapter, more specific conclusions from this thesis are provided in the first section. Then, by stating the shortcomings and limitations, some possible directions to extend the thesis are discussed.

7.1 Conclusions

This thesis was devoted to dealing with two practical challenges in modern wireless communications, i.e., the delays in CSIT as well as random node distribution. On the one hand, the available CSIT is not instantaneous because of the CSI feedback mechanism, but it is required in most of the beamforming techniques. On the other hand, randomness and the scale of wireless networks make it unreasonable to ignore the path-loss effect. To tackle these two challenges, two research approaches were followed: 1) The MAT scheme, which was the major research direction that utilised the delayed CSIT, and therefore it was meaningful to generalise this method and enhance its performance; 2) Stochastic geometry-based approaches, which have been exploited as a powerful tool to characterise and analyse the signal power in large-scale randomly distributed wireless networks.

Based on information-theoretical analysis, the MAT scheme was generalized and its performance

was enhanced using a probabilistic-constrained optimisation approach. The logic behind this was that, for a good communication system, it was unlikely that the transmitters and receivers would spend all of the DoF to cancel interference assuming an infinite SNR. Therefore, taking unknown channel entries into consideration, constraints were proposed to guarantee high probability that the interference leakage power remained under a certain threshold, assuming the presence of generalised minimum interference leakage receivers. Furthermore, the detectability of the desired signal was maximised. The proposed iterative-free algorithm was exploited for the BC with two system configurations: *two* users with 2-antenna BS and *three* users with 2-antenna BS. It was shown that the allowed interference leakage power factor γ has minor influence on the achievable sum rate once it is kept under a certain threshold. In other words, there is minor sum rate degradation if a small-enough γ is selected. This observation led to the conclusion that effectively balancing between aligning the interference and enhancing the detectability of the desired signal can be approximated by conservatively constraining the interference leakage power. Finally, it was shown that the proposed algorithm outperforms the MAT scheme when the spatial correlation at the receiver side is large. When this spatial correlation decreases, the performance gain becomes insignificant.

As the assumption of completely independent channel variation is overly pessimistic, it was meaningful to extend the MAT scheme to the time-correlated channels. DoF regions as well as achievability schemes for three specific scenarios were studied: 1) 2-user and $2N$-antenna transmitter MIMO (N receiver antennas) IC with imperfect current and delayed CSIT; 2) K-user and KN-antenna BS MIMO (N receiver antennas) BC with imperfect current CSIT and perfect delayed CSIT; 3) 2-user and $2N$-antenna BS MIMO (N receiver antennas) BC with imperfect current CSIT and perfect delayed CSIT. Note that the MIMO considerations were straightforward extensions of the corresponding MISO settings, i.e., when $N = 1$. It is also mention worthy to highlight the following remarks:

- For Scenario 1), there is no benefit to increasing the quality of delayed CSIT once $\beta \geq \frac{1+2\alpha}{3}$.
- For Scenario 2), $N\frac{K(1-\alpha)}{1+\frac{1}{2}+\frac{1}{3}+\cdots\frac{1}{K}} + NK\alpha$ DoF are achieved for the symmetric DoF vertex. For the asymmetric DoF vertices, one of the receivers is guaranteed N DoF and each of the other receivers obtain $N\alpha$ DoF.
- For Scenario 3), a 3-dimensional DoF region was proposed and proven to be convex. Symmetrically, each user achieves $N\frac{2+\alpha}{6}$ DoF. For asymmetric DoF vertices, if one of the receivers is guaranteed N DoF, the other two receivers obtain $N\alpha$ DoF in total.

The most important observation from the DoF regions was that the CSIT is always useful except for Scenario 1) when $\beta \geq \frac{1+2\alpha}{3}$. More specifically, when $\alpha = \beta = 0$, the receivers achieve N DoF in total, when $\alpha = 0$, $\beta = 1$, the DoF regions tend to be N times the ones achieved by the MAT scheme and when $\alpha = \beta = 1$, the DoF regions converge to the one achieved using the ZF receiver with perfect current CSIT. Finally, in Scenario 1), the achievability schemes can be extended from the IC to the BC and vice versa. In the other two scenarios, the schemes for the BC cannot be extended to the IC because special combinations of private symbols are required to be sent.

The performance of the ZF receivers was evaluated for multiple-antenna PPP wireless networks. A typical receiver was studied in details which represented the average performance of all the receivers. At first, bounds of transmission capacity were derived where the ZF receiver can design its CSI area and cancel all the inside interferers. It was shown that, in the case of large transmitter intensity, more SRDoF should be allocated to cancel interference. Afterwards, the influence of the feasibility condition of the ZF receiver to the transmission capacity was analysed. Finally, a dynamic wireless network was considered in which transmitters and receivers joined and left the network independently. The dynamic data traffic was modelled by queueing theory. The feasibility condition of the ZF receiver was studied, which determined the probability that it was allowed to join the network, i.e., the access probability. When increasing the number of data streams, the access probability decreases dramatically in a high intensity scenario. Meanwhile, the IA beamforming approach was utilised sequentially and the evolving success probability at the receiver was characterised based on the queueing model. Finally, the closed-form expression of the transmission capacity was derived, which illustrated that in order to achieve the highest area spectrum efficiency, only *one* data stream should be transmitted.

To make use of the distance information in randomly distributed wireless networks, a distance-dependent IA approach was designed for a generic 2-tier heterogeneous wireless network. The average throughput achieved by the second-tier receivers in a cluster was derived, which led to the optimal number of data streams from the nearest cross-tier transmitter that should be treated as interference. Essentially, throughput is maximised by considering a trade-off between increasing the number of data streams and enhancing the SIR at each data stream. The throughputs with accurate and statistical distance knowledge to the nearest cross-tier transmitter were researched. Interestingly, obtaining accurate distance information is of minor importance. It could also be observed that minor CSI error does not dramatically reduce the throughput using the proposed distance-dependent IA approach. After mitigating the nearest cross-tier interference, a stochastic

geometry technique was exploited to model the remaining cross-tier interference and the inter-cluster interference. Specifically, a near closed-form expression and a corresponding upper bound were derived for the success probability which was effected by the marginalised remaining cross-tier interference. Also, an approximation was introduced for the success probability that was affected by the inter-cluster interference, assuming a small cluster size. As a conclusion, the interference power generated by a Poisson clustered wireless network can be reduced by shrinking the cluster size.

7.2 Limitations and Future Work

Even though this thesis achieved many interesting results, some constraints limited the application of these results practically. The remainder of this section contains some interesting future research directions that can help to make our work more general and practically appealing.

7.2.1 Further Exploit the Delayed CSIT

Extending the MAT scheme to the time-correlated channels has been agreed as a successful way to exploit both the current and delayed CSIT. However, most of the work, such as [66] and [59], considered either the 2-user cases or certain antenna settings. Therefore, there is still a long way to go to achieve the DoF region and achievability schemes for the most generic scenario where a general number of users and antennas as well as the evolving quality of CSIT are considered.

Although the information-theoretical studies on the delayed CSIT have made a number of achievements, few have incorporated these results with optimisation tools to further enhance the achievable sum rate in the finite SNR case. Although a robust optimisation method was proposed in this thesis, just like the dual SINR algorithm in [46], it is only suitable for the time-independent channel. How to introduce the optimisation tools to enhance the achievable sum rate in the finite SNR case using both the current CSIT and the delayed CSIT is of great interest.

The achievability schemes for more than 2-user BC proposed in Section 4.5 cannot be extended to IC. This is because the BS is required to send special combinations of all the private symbols and this cannot be accomplished in the IC where the transmitters only know their local private symbols. Therefore, it is challenging and meaningful to extend the MAT scheme for more than 2-user cases to the IC. In other words, how to exploit the delayed CSIT in IC is of interest.

7.2.2 Realise Full Potential of Stochastic Geometry

More Practical System Model

Most of the work on stochastic geometry technique focus on the single-antenna system, homogeneous transmit power and independent node distribution, etc. However, these assumptions are far too ideal in the real world. For example, most BSs are equipped with more than one antennas and power allocated at the BSs are usually different. Therefore, considerations of more advanced system models including MIMO, spectrum allocation, mobility and uplink transmission are required.

This thesis assumes exponentially distributed desired signal power at each data stream, which represents the average behaviour. Specifically, if the transmitter has CSIT and designs the beamforming matrix using maximum ratio transmission (MRT), the received desired signal power will be related to the eigenvalue of the channel. Therefore, more complex analysis based on a combination of random matrix theorem and stochastic geometry should be made to accurately characterise the desired signal power. In this way, the optimal solution to achieve the highest transmission capacity may be different from that in Chapter 5. The authors in [9] have already begun this research direction by assuming a constant eigenvalue.

The combination of the queueing theory and the stochastic geometry technique is highlighted by [95]. Considering multiple access control (MAC), especially the carrier sense multiple access (CSMA) protocol and the buffer structure of network elements, transmission is thought to be bursty and concurrent transmission is not realistic. Moreover, the packet delay is an important metric for the performance of wireless networks, which can be characterised using the queuing theory. As a result, this combination of the queueing theory and the stochastic geometry-based analysis is practically appealing and more practical models are required to bring a better understanding of the system performance.

More Advanced Point Process

Compared to the lattice model, stochastic geometry-based approaches have remarkably enhanced the accuracy when modeling the wireless networks. The most fundamental model in stochastic geometry is PPP, which is overly ideal by assuming each node is totally independent with other nodes. Even though clustered point processes offset the deficiency of PPP by considering the

correlations between nodes, it is still far from a practically appealing model. Hard-core point process (HCPP) and Poisson hole process provide more realistic models by considering more sophisticated network planning and interference management techniques. However, finding a tractable closed-form expression for the node distribution is a challenge.

7.2.3 Other Future Directions

Evaluating the performance of the MAT-based schemes in the presence of path-loss effect and random node distribution is meaningful. Specifically, as discussed in [88], the quality of CSIT is a function of the distance between the transmitter-receiver pair and the training period. If Poisson distributed receivers are considered, qualities of CSIT from a BS to each receiver become a function of receiver distribution. In this way, the performance of these schemes can be characterised in the randomly distributed wireless network. Moreover, the minimum training period can also be obtained to guarantee a certain sum rate.

Note that the authors in [88] and [44] studied the transmission capacity achieved by diversity combining. Conversely, no diversity combination is considered in [9] and this thesis. Therefore, an interesting topic would be to analysis the multiplexing-diversity trade-off in distributed wireless networks such as Poisson networks. Transmission capacity is an ideal metric because it takes the outage probability and the multiplexing gain into consideration.

Appendix A

Proof of Theorem 1

The achievable sum rate for the e-channel should be written as,

$$I(\mathbf{s}_i;\mathbf{y}_i) = \log\det\left(\mathbf{I}^{(2)} + \frac{P}{2}\mathbf{N}_i^{-1}\underbrace{\left(\mathbf{I}+\frac{P}{2}\mathbf{G}_{ij}\mathbf{G}_{ij}^H\right)^{-1}\mathbf{G}_{ii}\mathbf{G}_{ii}^H}_{\bar{\mathbf{G}}_i}\right), i\neq j, i,j=1,2.$$

where $\mathbf{I}^{(2)}$ is the 2×2 identity matrix and $\mathbf{N}_i = \mathrm{diag}\{1, \max\{1,|g_i|^2\}\}$ is the covariance matrix of the white noise at Rx-i, where $\max\{1,|g_i|^2\}$ is denoted as $\hat{g}_i^2$. In the case of Rx-1, with $\mathbf{G}_{11}$ and $\mathbf{G}_{12}$ defined in (3.1), the following equation can be obtained:

$$\bar{\mathbf{G}}_1 = \begin{bmatrix} \|\mathbf{h}_1(1)\|^2 & h_{1,1}^*(3)\mathbf{h}_1(1)\mathbf{w}_1^H \\ \frac{h_{1,1}(3)\mathbf{w}_1\mathbf{h}_1^H(1)}{\mathcal{Q}_1} & \frac{|h_{1,1}(3)|^2\|\mathbf{w}_1\|^2}{\mathcal{Q}_1} \end{bmatrix},$$

where,

$$\mathcal{Q}_1 = 1+\frac{P}{2}\|h_{1,1}(3)\mathbf{w}_2 - g_1\mathbf{h}_1(2)\|^2.$$

Hence, the equality inside the logarithmic function at Rx-1 can be expanded as,

$$\begin{aligned}\det\left(\mathbf{I}+\frac{P}{2}\mathbf{N}_1^{-1}\bar{\mathbf{G}}_1\right) =& \det\left(\begin{bmatrix} 1+\frac{P}{2}\|\mathbf{h}_1(1)\|^2 & \frac{P}{2}h_{1,1}^*(3)\mathbf{h}_1(1)\mathbf{w}_1^H \\ \frac{Ph_{A1}(3)\mathbf{w}_1\mathbf{h}_1^H(1)}{2\mathcal{Q}_1\hat{g}_1^2} & 1+\frac{P|h_{1,1}(3)|^2\|\mathbf{w}_1\|^2}{2\mathcal{Q}_1\hat{g}_1^2}\end{bmatrix}\right)\\ =& 1+\frac{P}{2}\left(\|\mathbf{h}_1(1)\|^2+\frac{|h_{A1}(3)|^2\|\mathbf{w}_1\|^2}{\mathcal{Q}_1\hat{g}_1^2}\right)\\ &+\frac{P^2}{4}\left(\frac{|h_{A1}(3)|^2\|\mathbf{h}_1(1)\|^2\|\mathbf{w}_1\|^2 - |h_{A1}(3)|^2\mathbf{w}_1\mathbf{h}_1(1)\mathbf{h}_1^H(1)\mathbf{w}_1^H}{\mathcal{Q}_1\hat{g}_1^2}\right)\end{aligned} \tag{A.1}$$

Since logarithm function is monotonically increasing, maximising (A.1) leads to the maximum achievable sum rate.

Because the SNR at the BS is P, the power constraint of the beamforming vectors is $\|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 \leq 2$. In order to achieve the maximum sum rate, the equality should always be valid, i.e., $\|\mathbf{w}_1\|^2 + \|\mathbf{w}_2\|^2 = 2$. Since it is assumed that all the channel entries are i.i.d., symmetric power allocation could be reasonably implemented to approximate the realistic situation, i.e., $\|\mathbf{w}_1\| = \|\mathbf{w}_2\| = 1$. Afterwards, the achievable rate at Rx-1 can be simplified as follows:

$$\begin{aligned}
I(\mathbf{s}_1;\mathbf{y}_1) &= \log\left(1+\frac{P}{2}\left(\|\mathbf{h}_1(1)\|^2+\frac{|h_{1,1}(3)|^2}{\mathcal{Q}_1\hat{g}_1^2}\right)\right.\\
&\left.+\frac{P^2}{4}\left(\frac{|h_{1,1}(3)|^2\|\mathbf{h}_1(1)\|^2-|h_{1,1}(3)|^2\mathbf{w}_1\mathbf{h}_1^H(1)\mathbf{h}_1(1)\mathbf{w}_1^H}{\mathcal{Q}_1\hat{g}_1^2}\right)\right)\\
&=\log\left(1+\frac{P}{2}\|\mathbf{h}_1(1)\|^2\right)\\
&\quad+\log\left(1+\frac{\frac{P}{2}|h_{1,1}(3)|^2+\frac{P^2}{4}|h_{1,1}(3)|^2\left(\|\mathbf{h}_1(1)\|^2-\mathbf{w}_1\mathbf{h}_1^H(1)\mathbf{h}_1(1)\mathbf{w}_1^H\right)}{\left(1+\frac{P}{2}\|\mathbf{h}_1(1)\|^2\right)\mathcal{Q}_1\hat{g}_1^2}\right)
\end{aligned}$$

Similarly, for Rx-2, the achievable rate is given by,

$$\begin{aligned}
I(\mathbf{s}_2;\mathbf{y}_2) =&\log\left(1+\frac{P}{2}\|\mathbf{h}_2(2)\|^2\right)\\
&+\log\left(1+\frac{\frac{P}{2}|h_{2,1}(3)|^2+\frac{P^2}{4}|h_{2,1}(3)|^2\left(\|\mathbf{h}_2(2)\|^2-\mathbf{w}_2\mathbf{h}_2^H(2)\mathbf{h}_2(2)\mathbf{w}_2^H\right)}{\left(1+\frac{P}{2}\|\mathbf{h}_2(2)\|^2\right)\mathcal{Q}_2\hat{g}_2^2}\right),
\end{aligned}$$

where $\mathcal{Q}_2 = 1+\frac{P}{2}\|h_{2,1}(3)\mathbf{w}_1 - x_2\mathbf{h}_2(1)\|^2$. Hence, the achievable sum rate of the system is given by,

$$\begin{aligned}
&I(\mathbf{s}_1;\mathbf{y}_1)+I(\mathbf{s}_2;\mathbf{y}_2) \approx \log\left(1+\frac{P}{2}\|\mathbf{h}_1(1)\|^2\right)+\log\left(1+\frac{P}{2}\|\mathbf{h}_2(2)\|^2\right)\\
&\quad+\log\left(\frac{\frac{P}{2}|h_{1,1}(3)|^2+\frac{P^2}{4}|h_{1,1}(3)|^2\left(\|\mathbf{h}_1(1)\|^2-\mathbf{w}_1\mathbf{h}_1^H(1)\mathbf{h}_1(1)\mathbf{w}_1^H\right)}{\left(1+\frac{P}{2}\|\mathbf{h}_1(1)\|^2\right)\mathcal{Q}_1\hat{g}_1^2}\right)\\
&\quad+\log\left(\frac{\frac{P}{2}|h_{2,1}(3)|^2+\frac{P^2}{4}|h_{2,1}(3)|^2\left(\|\mathbf{h}_2(2)\|^2-\mathbf{w}_2\mathbf{h}_2^H(2)\mathbf{h}_1(1)\mathbf{w}_2^H\right)}{\left(1+\frac{P}{2}\|\mathbf{h}_2(2)\|^2\right)\mathcal{Q}_2\hat{g}_2^2}\right)\\
&=\log\left(1+\frac{P}{2}\|\mathbf{h}_1(1)\|^2\right)+\log\left(1+\frac{P}{2}\|\mathbf{h}_2(2)\|^2\right)\\
&\quad+\log\left(\frac{\frac{P}{2}|h_{1,1}(3)|^2+\frac{P^2}{4}|h_{1,1}(3)|^2\left(\|\mathbf{h}_1(1)\|^2-\mathbf{w}_1\mathbf{h}_1^H(1)\mathbf{h}_1(1)\mathbf{w}_1^H\right)}{\left(1+\frac{P}{2}\|\mathbf{h}_1(1)\|^2\right)\mathcal{Q}_2\hat{g}_2^2}\right) \qquad \text{(A.2a)}\\
&\quad+\log\left(\frac{\frac{P}{2}|h_{2,1}(3)|^2+\frac{P^2}{4}|h_{2,1}(3)|^2\left(\|\mathbf{h}_2(2)\|^2-\mathbf{w}_2\mathbf{h}_2^H(2)\mathbf{h}_2(2)\mathbf{w}_2^H\right)}{\left(1+\frac{P}{2}\|\mathbf{h}_2(2)\|^2\right)\mathcal{Q}_1\hat{g}_1^2}\right).
\end{aligned}$$

Note that (A.2a) is based on the property of logarithmic function.

Appendix B

Proof of Theorem 2

By rearranging the probabilistic constraint (3.6a), the following equation can be obtained:

$$\Pr\left(|h_{2,1}(3)|^2 \leq \frac{2\hat{\alpha}}{P\,\mathrm{trace}\left(\mathbf{w}_1\left(\mathbf{h}_2^H(1)\right)^{\perp}\mathbf{h}_2^{\perp}(1)\mathbf{w}_1^H\right)}\right) \geq \hat{\beta}.$$

By assuming $\mathcal{Q}_3 = \frac{2\hat{\alpha}}{P\,\mathrm{trace}\left(\mathbf{w}_1\left(\mathbf{h}_2^H(1)\right)^{\perp}\mathbf{h}_2^{\perp}(1)\mathbf{w}_1^H\right)}$, the previous expression can be simplified into the following equation:

$$\Pr\left(|h_{2,1}(3)| \leq \sqrt{\mathcal{Q}_3}\right) \geq \hat{\beta}. \tag{B.1}$$

In this inequality, knowledge of all the parameters are available at the BS except for $h_{2,1}(3)$ and with the distribution of $|h_{2,1}(3)|$, the constraint is equivalent to finding the cumulative distribution function (CDF) of $|h_{A1}(3)|$. Typically, the real and imaginary part of the channel are i.i.d., then,

$$\mathrm{Re}\{h_{A1}(3)\} \sim \mathcal{N}\left(0, \sigma_h^2\right);\ \mathrm{Im}\{h_{A1}(3)\} \sim \mathcal{N}\left(0, \sigma_h^2\right)$$

where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ are the real and imaginary parts, respectively, and σ_h^2 is the variance of channel entries. Therefore, the probabilistic constraint can be approximated as follows:

$$\begin{aligned}
&\Pr\left(|h_{2,1}(3)| \leq \sqrt{\mathcal{Q}_3}\right) \\
\geq &\Pr\left(|\,\mathrm{Re}\{h_{2,1}(3)\}| \leq \sqrt{\frac{\mathcal{Q}_3}{2}}\right) \bigcap \Pr\left(|\,\mathrm{Im}\{h_{2,1}(3)\}| \leq \sqrt{\frac{\mathcal{Q}_3}{2}}\right) \\
= &\Pr\left(|\,\mathrm{Re}\{h_{2,1}(3)\}| \leq \sqrt{\frac{\mathcal{Q}_3}{2}}\right) \cdot \Pr\left(|\,\mathrm{Im}\{h_{2,1}(3)\}| \leq \sqrt{\frac{\mathcal{Q}_3}{2}}\right) \\
= &\left(\Pr\left(|\,\mathrm{Re}\{h_{2,1}(3)\}| \leq \sqrt{\frac{\mathcal{Q}_3}{2}}\right)\right)^2 = \left(\Pr\left(|\,\mathrm{Im}\{h_{2,1}(3)\}| \leq \sqrt{\frac{\mathcal{Q}_3}{2}}\right)\right)^2.
\end{aligned} \tag{B.2}$$

With (B.1) and (B.2), the probabilistic constraint can be further written as,

$$\Pr\left(|\mathrm{Re}\{h_{2,1}(3)\}| \leq \sqrt{\frac{\mathcal{Q}_3}{2}}\right) \geq \sqrt{\hat{\beta}},\ \text{or equivalently,}\ \Pr\left(|\mathrm{Im}\{h_{2,1}(3)\}| \leq \sqrt{\frac{\mathcal{Q}_3}{2}}\right) \geq \sqrt{\hat{\beta}},$$

In the rest of the derivation, only the probabilistic constraint with the real part of $h_{2,1}(3)$ is considered. As the real and imaginary part of the channel entries are i.i.d., exactly the same constraint should be obtained if the imaginary part is considered. Using the CDF of Gaussian distribution with *zero* mean and variance σ_h^2, the probabilistic constraint can be further rewritten as,

$$\begin{aligned}\Pr\left(|\operatorname{Re}\{h_{2,1}(3)\}| \leq \sqrt{\frac{\mathcal{Q}_3}{2}}\right) =& \Pr\left(\operatorname{Re}\{h_{2,1}(3)\} \leq \sqrt{\frac{\mathcal{Q}_3}{2}}\right) - \Pr\left(\operatorname{Re}\{h_{2,1}(3)\} \leq -\sqrt{\frac{\mathcal{Q}_3}{2}}\right) \\ =& \frac{1}{2}\left(\operatorname{erf}\left(\frac{\sqrt{\mathcal{Q}_3}}{2\sigma_h}\right) - \operatorname{erf}\left(\frac{-\sqrt{\mathcal{Q}_3}}{2\sigma_h}\right)\right), \end{aligned} \tag{B.3}$$

where $\operatorname{erf}(x)$ is the standard error function. Afterwards, based on the property of the error function,

$$\operatorname{erf}\left(\frac{\sqrt{\mathcal{Q}_3}}{2\sigma_h}\right) \geq \sqrt{\hat{\beta}} \quad \Longrightarrow \quad \mathcal{Q}_3 \geq 4\sigma_h^2 \left(\operatorname{erf}^{-1}(\sqrt{\hat{\beta}})\right)^2.$$

Finally, by expanding $\mathcal{Q}_3$, the probabilistic constraint can be written as:

$$\operatorname{trace}\left(\mathbf{w}_1 \left(\mathbf{h}_2^H(1)\right)^{\perp} \mathbf{h}_2^{\perp}(1)\mathbf{w}_1^H\right) \leq \frac{\hat{\alpha}}{2\sigma_h^2 P\left(\operatorname{erf}^{-1}(\sqrt{\hat{\beta}})\right)^2}.$$

Appendix C

Proof of Theorem 8

Firstly, because of the independence between the fading effect and the path-loss effect, $\mathcal{Q}_6$ can be written as the k-th power of the outage probability affected by the interference from one transmitter, i.e.,

$$\begin{aligned}\mathcal{Q}_6 &= \underbrace{\int_{r_z}^{r_\infty}\cdots\int_{r_z}^{r_\infty}}_{k}\int_0^\infty \exp\left(-\frac{\theta x l_{00}^{\varsigma}\boldsymbol{\omega}^{-\varsigma}}{2}\right) f_h(x) f_\Omega(\boldsymbol{\omega})\mathrm{d}x\mathrm{d}\omega_1\mathrm{d}\omega_2\cdots\mathrm{d}\omega_k \\ &= \prod_k \int_{r_z}^{r_\infty}\int_0^\infty \exp\left(-\frac{\theta x l_{00}^{\varsigma}\omega^{-\varsigma}}{2}\right) f_h(x) f_\Omega(\omega)\mathrm{d}x\mathrm{d}\omega \\ &= \left(\int_{r_z}^{r_\infty}\int_0^\infty \exp\left(-\frac{\theta x l_{00}^{\varsigma}\omega^{-\varsigma}}{2}\right) f_h(x) f_\Omega(\omega)\mathrm{d}x\mathrm{d}\omega\right)^k.\end{aligned}$$

Taking the expression from (5.16) and (5.17) into $\mathcal{Q}_6$, the following equation can be obtained:

$$\begin{aligned}\mathcal{Q}_6 &= \left(\int_{r_z}^{r_\infty}\int_0^\infty \exp\left(-\frac{\theta x l_{00}^{\varsigma}\omega^{-\varsigma}}{2}\right)\frac{x^{\bar{d}-1}e^{-x/2}}{2^d\Gamma(\bar{d})}\frac{2\omega}{r_\infty^2-r_z^2}\mathrm{d}x\mathrm{d}\omega\right)^k \\ &= \left(\int_{r_z}^{r_\infty}\frac{\omega(\theta l_{00}^{\varsigma}\omega^{-\varsigma}/2+1/2)^{-\bar{d}}}{(r_\infty^2-r_z^2)2^{\bar{d}-1}}\mathrm{d}\omega\right)^k.\end{aligned}$$

Note that $\mathcal{Q}_6$ is the success probability if there is k transmitters on the plane outside r_z. Taking the Poisson distribution into consideration, the averages success probability is given by,

$$\begin{aligned}p_s &= \sum_{k=0}^{\infty}\frac{\exp(-\pi\lambda(r_\infty^2-r_z^2))(\pi\lambda(r_\infty^2-r_z^2))^k}{k!}\mathcal{Q}_k \\ &= \exp\left(\pi\lambda(r_\infty^2-r_z^2)\left(\int_{r_z}^{r_\infty}\frac{\omega(\theta l_{00}^{\varsigma}\omega^{-\varsigma}/2+1/2)^{-\bar{d}}}{(r_\infty^2-r_z^2)2^{\bar{d}-1}}d\omega-1\right)\right) \\ &= \exp\left(\frac{2^{\bar{d}}r_z^2-(\theta l_{00}^{\varsigma}r_z^{-\varsigma}/2+1/2)^{-\bar{d}}r_z^2-\frac{\varsigma\theta\bar{d}}{2}\mathcal{Q}_7}{(\pi\lambda)^{-1}2^{\bar{d}}}\right),\end{aligned} \tag{C.1a}$$

where,

$$\begin{aligned}\mathcal{Q}_7 &= \int_{r_\infty}^{\infty} \omega^2 (\theta l_{00}^{\varsigma} \omega^{-\varsigma}/2 + 1/2)^{-\bar{d}-1} \omega^{-\varsigma-1} \mathrm{d}\omega \\ &= \left(2^{1+\bar{d}} r_z^{2-\varsigma}\right) \frac{{}_2F_1\left[\frac{\varsigma-2}{\varsigma}, 1+\bar{d}, 2-\frac{2}{\varsigma}, -\theta l_{00}^{\varsigma} r_z^{-\varsigma}\right]}{\varsigma - 2}, \end{aligned} \tag{C.2a}$$

where (C.1a) and (C.2a) is because transmitters are distributed in an infinite plane, i.e., $r_\infty \to \infty$.

Appendix D

Proof of Theorem 9

The interference power from L_0 at the q-th data stream of R_0 is given by

$$P_A l_{0L_0}^{-\varsigma} \left\| \mathbf{f}_q^H \mathbf{H}_{iL_0} \mathbf{V}_{L_0} \right\|^2,$$

Because each of the channel entry is complex Gaussian distribution with unit-variance and zero-mean, $\left\| \mathbf{f}_q^H \mathbf{H}_{iL_0} \mathbf{V}_{L_0} \right\|^2$ is chi-squared distributed with $2d$ DoF with the PDF given by (5.16). Therefore, the success probability can be expressed as,

$$\begin{aligned} p_s^n &= \int_0^\infty \exp\left(-Pl^{-\varsigma}\theta l_{00}^{\varsigma} x\right) f_{\chi^2}(2x, 2d)\mathrm{d}2x \\ &= \int_0^\infty \exp\left(-P_A l_{0L_0}^{-\varsigma}\theta l_{00}^{\varsigma} x\right) \frac{2(2x)^{d-1}e^{-x}}{2^d\Gamma(d)}\mathrm{d}x \\ &= \frac{\int_0^\infty 2(2x)^{d-1}\exp\left(-P_A l_{0L_0}^{-\varsigma}\theta l_{00}^{\varsigma} x - x\right)\mathrm{d}x}{2^d\Gamma(d)} = \left(P_A l_{0L_0}^{-\varsigma}\theta l_{00}^{\varsigma} + 1\right)^{-d}. \end{aligned}$$

Appendix E

Proof of Theorem 10

Taking the PDF of the distance to the nearest point in PPP $f_{L_0}(x)$ into (6.10), the following equation can be obtained:

$$p_s^A = \exp\left(-(2\pi\lambda_A)^2\mathbb{E}_h\left[\int_0^\infty\int_x^\infty\left(1-\exp\left(-\frac{\theta l_{00}^\varsigma h P_A l^{-\varsigma}}{2}\right)\right) lx\exp(-\pi\lambda_A x^2)\mathrm{d}l\mathrm{d}x\right]\right).$$

With the following integral,

$$\int_x^\infty\left(1-\exp\left(-\frac{\theta l_{00}^\varsigma h P_A l^{-\varsigma}}{2}\right)\right) l\mathrm{d}l = -\frac{x^2}{2}+\frac{1}{\varsigma}\left(\frac{h\theta l_{00}^\varsigma P_A}{2}\right)^{2/\varsigma}\times \left(-\Gamma\left(-\frac{2}{\varsigma}\right)+\Gamma\left(-\frac{2}{\varsigma},\frac{h\theta l_{00}^\varsigma P_A x^{-\varsigma}}{2}\right)\right),$$

and,

$$\int_0^\infty -\frac{x^2}{2}\left[x\exp\left(-\pi\lambda_A x^2\right)\right]\mathrm{d}x = -\frac{1}{4\pi^2\lambda_A^2}.$$

The success probability can be reduced to the following equation:

$$\begin{aligned} p_s^A =& \exp\left(\mathbb{E}_h\left[1+\frac{2\pi\lambda_A\Gamma\left(-\frac{2}{\varsigma}\right)}{\varsigma}\left(\frac{h\theta l_{00}^\varsigma P_A}{2}\right)^{2/\varsigma}\right.\right. \\ &\left.\left.-\frac{(2\pi\lambda)^2}{\varsigma}\left(\frac{h\theta l_{00}^\varsigma P_A}{2}\right)^{2/\varsigma}\int_0^\infty x\exp(-\pi\lambda_A x^2)\Gamma\left(-\frac{2}{\varsigma},\frac{h\theta l_{00}^\varsigma P_A x^{-\varsigma}}{2}\right)\mathrm{d}x\right]\right). \end{aligned} \tag{E.1}$$

The expectation operation in terms of the fading coefficient h can be simply solved for the first two terms. However, the third term needs further derivation, i.e.,

$$\begin{aligned} &\int_0^\infty x\exp(-\pi\lambda_A x^2)\Gamma\left(-\frac{2}{\varsigma},\frac{h\theta l_{00}^\varsigma P_A x^{-\varsigma}}{2}\right)\mathrm{d}x \\ &= -\frac{1}{2\pi\lambda}\int_0^\infty\Gamma\left(-\frac{2}{\varsigma},\frac{h\theta l_{00}^\varsigma P_A x^{-\varsigma}}{2}\right)\mathrm{d}\exp(-\pi\lambda_A x^2) \end{aligned}$$

$$= -\frac{1}{2\pi\lambda_A}\left(\Gamma\left(-\frac{2}{\varsigma}, \frac{h\theta l_{00}^{\varsigma} P_A x^{-\varsigma}}{2}\right)\exp(-\pi\lambda_A x^2)\Big|_0^{\infty} - \int_0^{\infty}\exp(-\pi\lambda_A x^2)\mathrm{d}\Gamma\left(-\frac{2}{\varsigma}, \frac{h\theta l_{00}^{\varsigma} P_A x^{-\varsigma}}{2}\right)\right) \tag{E.2a}$$

$$= \frac{1}{2\pi\lambda_A}\int_0^{\infty}\exp(-\pi\lambda_A x^2)\mathrm{d}\Gamma\left(-\frac{2}{\varsigma}, \frac{h\theta l_{00}^{\varsigma} P_A x^{-\varsigma}}{2}\right) \tag{E.2b}$$

$$= \frac{1}{2\pi\lambda_A}\int_0^{\infty}\exp(-\pi\lambda_A x^2)\mathrm{d}\int_{\frac{h\theta l_{00}^{\varsigma} P_A x^{-\varsigma}}{2}}^{\infty} t^{-\frac{2}{\varsigma}-1}e^{-t}\mathrm{d}t$$

$$= \frac{1}{2\pi\lambda_A}\int_0^{\infty}\exp(-\pi\lambda_A x^2)\frac{\mathrm{d}\frac{h\theta l_{00}^{\varsigma} P_A x^{-\varsigma}}{2}}{\mathrm{d}x}\left(\frac{h\theta l_{00}^{\varsigma} P_A x^{-\varsigma}}{2}\right)^{-\frac{2}{\varsigma}-1}\exp\left(-\frac{h\theta l_{00}^{\varsigma} P_A x^{-\varsigma}}{2}\right)\mathrm{d}x \tag{E.2c}$$

$$= \frac{\varsigma}{4\pi\lambda_A}\left(\frac{h\theta l_{00}^{\varsigma} P_A x^{-\varsigma}}{2}\right)^{-\frac{2}{\varsigma}}\int_0^{\infty}\exp\left(-\pi\lambda_A t - \frac{h\theta l_{00}^{\varsigma} P_A t^{-\varsigma/2}}{2}\right)\mathrm{d}t,$$

where (E.2a) is through integration by parts. Equation (E.2b) is based on the following equations:

$$\lim_{x\to\infty}\Gamma\left(-\frac{2}{\varsigma}, \frac{h\theta l_{00}^{\varsigma} P_A x^{-\varsigma}}{2}\right) = \Gamma\left(-\frac{2}{\varsigma}\right), \text{because } \varsigma > 2, \text{it is a real number,}$$

$$\lim_{x\to\infty}\exp(-\pi\lambda_A x^2) = 0,$$

$$\lim_{x\to 0}\Gamma\left(-\frac{2}{\varsigma}, \frac{h\theta l_{00}^{\varsigma} P_A x^{-\varsigma}}{2}\right) = 0,$$

$$\lim_{x\to 0}\exp(-\pi\lambda_A x^2) = 1,$$

and (E.2c) is according to Newton-Leibniz axiom. Because the PDF of the fading coefficient is chi-squared distributed with $2\bar{d}_A$ DoF, the third term of (E.1) can be translated using the following equations:

$$\mathbb{E}_h\left[\frac{(2\pi\lambda_A)^2}{\varsigma}\left(\frac{h\theta l_{00}^{\varsigma} P_A}{2}\right)^{2/\varsigma}\frac{\varsigma\left(\frac{h\theta l_{00}^{\varsigma} P_A x^{-\varsigma}}{2}\right)^{-\frac{2}{\varsigma}}}{4\pi\lambda_A}\int_0^{\infty}\exp\left(-\pi\lambda_A t - \frac{h\theta l_{00}^{\varsigma} P_A t^{-\varsigma/2}}{2}\right)\mathrm{d}t\right]$$

$$= \mathbb{E}_h\left[\pi\lambda_A\int_0^{\infty}\exp\left(-\pi\lambda_A t - \frac{h\theta l_{00}^{\varsigma} P_A t^{-\varsigma/2}}{2}\right)\mathrm{d}t\right]$$

$$= \pi\lambda_A\int_0^{\infty}\int_0^{\infty}\exp\left(-\pi\lambda_A t - \frac{h\theta l_{00}^{\varsigma} P_A t^{-\varsigma/2}}{2}\right)\frac{x^{\bar{d}_A-1}e^{-x/2}}{2^{\bar{d}_A}\Gamma(\bar{d}_A)}\mathrm{d}x\mathrm{d}t$$

$$= \pi\lambda_A\int_0^{\infty} e^{-\pi\lambda_A t}\left(1 + P_A\theta l_{00}^{\varsigma} t^{-\varsigma/2}\right)^{-\bar{d}_A}\mathrm{d}t.$$

Because the $\frac{2}{\varsigma}$-th moment of h can be written as,

$$\mathbb{E}\left[h^{2/\varsigma}\right] = \frac{2^{2/\varsigma}\Gamma(\bar{d}_A + 2/\varsigma)}{\Gamma(\bar{d}_A)}.$$

Hence, the success probability can be finalised as,

$$p_s^A = \exp\left(1 + \frac{2\pi\lambda_A(P_A\theta l_{00}^{\varsigma})^{\frac{2}{\varsigma}}\Gamma(\bar{d}_A + \frac{2}{\varsigma})\Gamma(-\frac{2}{\varsigma})}{\varsigma\Gamma(\bar{d}_A)} - \pi\lambda_A\int_0^{\infty} e^{-\pi\lambda_A t}(1 + P_A\theta l_{00}^{\varsigma}t^{-\varsigma/2})^{-\bar{d}_A}\mathrm{d}t\right).$$

Appendix F

List of Publications

This Appendix contains a list of published and submitted papers.

F.1 Accepted Publications

- **Yi Luo** and T. Ratnarajah, "Robust Stochastic Optimization For MISO Broadcasting Channel With Delayed CSIT and Limited Transmitting Antennas," Accepted by *IEEE Trans. Veh. Technol.*, Sep, 2014.
- **Yi Luo**, H. Du, J. Xue, T. Ratnarajah and D. Wilcox, "On The Performance of Asynchronous Ad-Hoc Networks Using Interference Alignment with Renewal Process," *IET Communications* , vol.7, no.15, pp.1628-1637, October 15 2013.
- **Yi Luo**, H. Du, T. Ratnarajah and D. Wilcox, "Performance of Homogeneous and Asynchronous Ad-Hoc Network with Interference Alignment," *In Proc. the IEEE International Conference on Communications (ICC)*, pp.1805-1809, June 9-13, 2013.
- **Yi Luo**, H. Du, T. Ratnarajah and D. Wilcox, "Performance Analysis of Ad-Hoc Networks with Interference Alignment," *In Proc. the IEEE Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pp.1412-1416, November 4-7, 2012.
- **Yi Luo**, A. Papazafeiropoulos and T. Ratnarajah, "Degrees-of-freedom Region of MIMO Interference Channel with Imperfect Current and Delayed CSIT," *In Proc. the IEEE Wireless Communications and Networking Conference (WCNC)*, pp.863-868, Istanbul, Turkey, April 6-9, 2014.
- **Yi Luo**, H. Du and T. Ratnarajah, "Robust Beamforming For Two-User MISO Broadcasting Channel with Delayed CSIT," *In Proc. on the Eighth IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp.161-164, A Coruna, Spain, June 22-25, 2014. [Nominated for student paper competition].

- **Yi Luo** and T. Ratnarajah, "Throughput of Two-Tier Heterogeneous Wireless Networks with Interference Coordination," *In Proc. the IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Washington, DC, USA, September 2-5, 2014.

F.2 Papers Under Revision

- **Yi Luo** and T. Ratnarajah, "Achieve Throughput of Two-Tier Randomly Distributed Heterogeneous Wireless Networks Using Interference Alignment,", *IEEE Trans. Comm.*, Revised on July 16, 2014.
- **Yi Luo**, T. Ratnarajah and A. Papazafeiropoulos, "Degrees-of-Freedom Regions for K-User MISO Time-Correlated Broadcast Channel,", Submitted to *IEEE International Conference on Communications (ICC)*, 2015.

Bibliography

[1] T. S. Rappaport. *Wireless communications principles and practices*. Prentice-Hall, 2002.

[2] D. Cavalcanti, D. Agrawal, Carlos Cordeiro, Bin Xie, and A. Kumar. Issues in integrating cellular networks WLANs, and MANETs: a futuristic heterogeneous wireless network. *IEEE Trans. Wireless Commun.*, 12(3):30–41, June 2005.

[3] M.A. Maddah-Ali, A.S. Motahari, and A.K. Khandani. Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis. *IEEE Trans. Inf. Theory*, 54(8):3457–3470, August 2008.

[4] S.A. Jafar and S. Shamai. Degrees of freedom region of the MIMO X channel. *IEEE Trans. Inf. Theory*, 54(1):151–170, January 2008.

[5] M. A. Maddad-Ali, A. S. Motahari, and A. K. Khandani. Signaling over MIMO multi-base systems: combination of multi-access and broadcast schemes. In *Proc. IEEE International Symposium on Information Theory, ISIT*, pages 2104–2108, July 2006.

[6] M.A. Maddah-Ali, A. Motahari, and A. Khandani. Communication over X channel: Signalling and multiplexing gain. Technical report, Univ. california, berkeley, CA, Tech. Rep., 2006.

[7] V.R. Cadambe and S.A. Jafar. Interference alignment and degrees of freedom of the K-user interference channel. *IEEE Trans. Inf. Theory*, 54(8):3425–3441, 2008.

[8] M.A. Maddah-Ali and D. Tse. Completely stale transmitter channel state information is still very useful. *IEEE Trans. Inf. Theory*, 58(7):4418–4431, 2012.

[9] R. Vaze and R.W. Heath. Transmission capacity of ad-hoc networks with multiple antennas using transmit stream adaptation and interference cancellation. *IEEE Trans. Inf. Theory*, 58(2):780–792, February 2012.

[10] S.P. Weber, X. Yang, J.G. Andrews, and G. de Veciana. Transmission capacity of wireless ad hoc networks with outage constraints. *IEEE Trans. Inf. Theory*, 51(12):4091 –4102, December 2005.

[11] S.W. Peters and R.W. Heath. Cooperative algorithms for mimo interference channels. *IEEE Trans. Veh. Technol.*, 60(1):206–218, January 2011.

[12] S. Zhang, S.C. Liew, and H. Wang. Blind known interference cancellation. *IEEE J. Sel. Areas Commun.*, 31(8):1572–1582, August 2013.

[13] X. Shang and H.V. Poor. Capacity region of vector gaussian interference channels with generally strong interference. *IEEE Trans. Inf. Theory*, 58(6):3472–3496, June 2012.

[14] V.S. Annapureddy and V.V. Veeravalli. Sum capacity of MIMO interference channels in the low interference regime. *IEEE Trans. Inf. Theory*, 57(5):2565–2581, May 2011.

[15] P. Tse, D.and Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press New York, NY, USA, 2005.

[16] K.T. Herring, J.W. Holloway, D.H. Staelin, and D.W. Bliss. Path-loss characteristics of urban wireless channels. *IEEE Trans. Antennas Propag.*, 58(1):171–177, January 2010.

[17] C. Xiao, Y.R. Zheng, and N.C. Beaulieu. Statistical simulation models for rayleigh and rician fading. In *Proc. IEEE International Conference on Communications (ICC)*, volume 5, pages 3524–3529, May 2003.

[18] R.K. Mallik. A new statistical model of the complex Nakagami-m fading gain. *IEEE Trans. Commun.*, 58(9):2611–2620, September 2010.

[19] H.Y. Shang, Y. Han, and J.H. Lu. Statistical analysis of Rician and Nakagami-m fading channel using multipath shape factors. In *Proc. Second International Conference on Computational Intelligence and Natural Computing Proceedings (CINC)*, volume 1, pages 398–401, September 2010.

[20] H. Wang, X. Zhou, and M.C. Reed. Physical layer security in cellular networks: A stochastic geometry approach. *IEEE Trans. Wireless Commun.*, 12(6):2776–2787, June 2013.

[21] R. Vaze, K.T. Truong, S. Weber, and R.W. Heath. Two-way transmission capacity of wireless ad-hoc networks. *IEEE Trans. Wireless Commun.*, 10(6):1966–1975, June 2011.

[22] R. Vaze. Transmission capacity of spectrum sharing ad hoc networks with multiple antennas. *IEEE Trans. Wireless Commun.*, 10(7):2334–2340, July 2011.

[23] S.P. Weber, J.G. Andrews, X. Yang, and G. Veciana. Transmission capacity of wireless ad hoc networks with successive interference cancellation. *IEEE Trans. Inf. Theory*, 53(8):2799–

2814, August 2007.

[24] L.H. Brandenburg and A.D. Wyner. Capacity of the gaussian channel with memory: The multivariate case. *Bell System Technical Journal*, 53:745–778, Brandenburg1974.

[25] J. Salz. Digital transmission over cross-coupled linear channels. *AT&T Technical Journal*, 64:1147–1159, 1985.

[26] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran. Multiuser MIMO achievable rates with downlink training and channelstate feedback. *IEEE Trans. Inf. Theory*, 56(6):2845–2866, 2010.

[27] X. Chen, S.H. Song, and K.B. Letaief. Interference alignment in MIMO interference relay channels. In *Proc. Wireless Communications and Networking Conference WCNC*, pages 630 –634, April 2012.

[28] C. Suh and D. Tse. Interference alignment for cellular networks. In *Communication , Control, and Computing*, pages 1037–1044, September 2008.

[29] M. Shen, C. Zhao, X. Liang, and Z. Ding. Best-effort interference alignment in OFDM systems with finite snr. In *Proc. IEEE International Conference on Communications (ICC)*, pages 1–6, June 2011.

[30] H. Ning, C. Ling, and K.K. Leung. Interference alignment with diversity. In *Sensor Signal Processing for Defence (SSPD)*, pages 1–5, September 2010.

[31] B. Nosrat-Makouei, J.G. Andrews, and R.W. Heath. MIMO interference alignment over correlated channels with imperfect CSI. *IEEE Trans. Signal Process.*, 59(6):2783–2794, 2011.

[32] H. Maleki, S.A. Jafar, and S. Shamai. Retrospective interference alignment over interference networks. *IEEE J. Sel. Topics Signal Process*, 6(3):228–240, June 2012.

[33] T. Xu and X.G. Xia. A diversity analysis for distributed interference alignment using the max-SINR algorithm. *IEEE Trans. Inf. Theory*, 60(3):1857–1868, March 2014.

[34] Y.F. Liu, Y.H. Dai, and Z.Q. Luo. On the complexity of leakage interference minimization for interference alignment. In *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 471–475, June 2011.

[35] C.M. Yetis, T. Gou, S.A. Jafar, and A.H. Kayran. On feasibility of interference alignment

in MIMO interference networks. *IEEE Trans. Signal Process.*, 58(9):4771–4782, September 2010.

[36] M. Haenggi, J.G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti. Stochastic geometry and random graphs for the analysis and design of wireless networks. *IEEE J. Sel. Areas Commun.*, 27(7):1029–1046, September 2009.

[37] Martin. Haenggi. *Stochastic Geometry for Wireless Networks*. Cambridge University Press, 2012.

[38] F. Baccelli and S. Zuyev. Stochastic geometry models of mobile communication networks. In *Frontiers in queueing: models and applications in science and engineering*, pages 227–243. CRC Press, 1996.

[39] M. Haenggi and R.K. Ganti. Interference in large wireless networks. *Foundations and Trends in Networking*, 3(2), February 2009.

[40] S. Landström, A. FuruSkär, J. Klas, L. Falconetti, and Kronestedt F. Heterogeneous networks increasing cellular capacity. *Ericsson review*, 1, 2011.

[41] Y. Kishiyama, A. Benjebbour, T. Nakamura, and H. Ishii. Future steps of LTE-A: evolution toward integration of local area and wide area systems. *IEEE Trans. Wireless Commun.*, 20(1):12–18, February 2013.

[42] J.G. Andrews, F. Baccelli, and R.K. Ganti. A tractable approach to coverage and rate in cellular networks. *IEEE Trans. Commun.*, 59(11):3122–3134, November 2011.

[43] F. Aurenhammer and H. Edelsbrunne. An optimal algorithm for constructing the weighted voronoi diagram in the panel. *Pattern Recognition*, 17(2):251 – 257, 1984.

[44] R.W. Heath and M. Kountouris. Modeling heterogeneous network interference with using poisson point processes. *IEEE Trans. Signal Process.*, 61(16):4114–4126, August 2013.

[45] X. Yi and D. Gesbert. Precoding methods for the MISO broadcast channel with delayed CSIT. *IEEE Trans. Wireless Commun.*, 12(5):1–11, 2013.

[46] Yi. X and D. Gesbert. Precoding on the broadcast MIMO channel with delayed CSIT: The finite SNR case. In *Proc IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP*, pages 2933–2936, 2012.

[47] S.A. Vorobyov, A.B. Gershman, and Z.-Q. Luo. Robust adaptive beamforming using worst-

case performance optimization: a solution to the signal mismatch problem. *IEEE Trans. Signal Process.*, 51(2):313–324, 2003.

[48] A. J G Anandkumar, A. Anandkumar, S. Lambotharan, and J.A. Chambers. Robust rate maximization game under bounded channel uncertainty. *IEEE Trans. Veh. Technol.*, 60(9):4471–4486, 2011.

[49] G. Zheng, K.K. Wong, and B. Ottersten. Robust cognitive beamforming with bounded channel uncertainties. *IEEE Trans. Signal Process.*, 57(12):4871–4881, December 2009.

[50] S.J. Kim, A. Magnani, A. Mutapcic, S.P. Boyd, and Z.Q. Luo. Robust beamforming via worst-case sinr maximization. *IEEE Trans. Signal Process.*, 56(4):1539–1547, April 2008.

[51] Y. Rong, S.A. Vorobyov, and A.B. Gershman. Robust linear receivers for multiaccess space-time block-coded mimo systems: a probabilistically constrained approach. *IEEE J. Sel. Areas Commun.*, 24(8):1560–1570, 2006.

[52] S.A. Vorobyov, Y. Rong, and A.B. Gershman. Robust adaptive beamforming using probability-constrained optimization. In *Proc. IEEE/SP 13th Workshop on Statistical Signal Processing*, pages 934–939, 2005.

[53] B.K. Chalise, S. Shahbazpanahi, A Czylwik, and AB. Gershman. Robust downlink beamforming based on outage probability specifications. *IEEE Trans. Wireless Commun.*, 6(10):3498–3503, October 2007.

[54] J. Chen and P. Elia. Degrees of freedom region of the MISO broadcast channel with general mixed-CSIT. *ArXiv e-prints*, May 2012.

[55] S. Yang, M. Kobayashi, P. Piantanida, and S. Shamai. Secrecy degrees of freedom of mimo broadcast channels with delayed csit. *IEEE Trans. Inf. Theory*, 59(9):5244–5256, 2013.

[56] Dimitri P. Bertsekas and Dimitri P. Bertsekas. *Nonlinear Programming*. Athena Scientific, 2nd edition, September 1999.

[57] A. Wiesel, Y. Eldar, and S. Shamai. Zero-forcing precoding and generalized inverses. *IEEE Trans. Signal Process.*, 55(9):4409–4418, September 2008.

[58] D. Shiu, G.J. Foschini, M.J. Gans, and J.M. Kahn. Fading correlation and its effect on the capacity of multielement antenna systems. *IEEE Trans. Commun.*, 48(3):502–513, March 2000.

[59] X. Yi, S. Yang, D. Gesbert, and M. Kobayashi. The degrees of freedom region of temporally correlated mimo networks with delayed CSIT. *IEEE Trans. Inf. Theory*, 60(1):494–514, January 2014.

[60] J. Chen and P. Elia. Can imperfect delayed CSIT be as useful as perfect delayed CSIT? dof analysis and constructions for the BC. In *50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1254–1261, 2012.

[61] J. Chen and P. Elia. Miso broadcast channel with delayed and evolving csit. In *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pages 992–996, July 2013.

[62] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi. Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT. *IEEE Trans. Inf. Theory*, 59(1):315–328, 2013.

[63] A. Papazafeiropoulos and T. Ratnarajah. Degrees of freedom of multiple-antenna interference channel withgeneral CSIT. In *IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pages 1108–1113, September 2013.

[64] X. Yi, D. Gesbert, S. Yang, and M. Kobayashi. On the DoF of the multiple-antenna time correlated interference channel with delayed CSIT. In *Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pages 1566–1570, November 2012.

[65] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. Wiley-Interscience, New York, NY, USA, 1991.

[66] J. Chen and P. Elia. Toward the performance versus feedback tradeoff for the two-user MISO broadcast channel. *IEEE Trans. Inf. Theory*, 59(12):8336–8356, December 2013.

[67] P. De Kerret, X. Yi, and D. Gesbert. On the degrees of freedom of the K-user time correlated broadcast channel with delayed CSIT. In *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pages 624–628, July 2013.

[68] P. Gupta and P.R. Kumar. The capacity of wireless networks. *IEEE Trans. Inf. Theory*, 46(2):388–404, March 2000.

[69] S. Toumpis and A.J. Goldsmith. Capacity regions for wireless ad hoc networks. *IEEE Trans. Wireless Commun.*, 2(4):736–748, July 2003.

[70] A. Zemlianov and Gustavo de Veciana. Capacity of ad hoc wireless networks with

infrastructure support. *IEEE J. Sel. Areas Commun.*, 23(3):657–667, March 2005.

[71] A.M. Hunter, J.G. Andrews, and S. Weber. Transmission capacity of ad hoc networks with spatial diversity. *IEEE Trans. Wireless Commun.*, 7(12):5058–5070, 2008.

[72] S. Weber, J.G. Andrews, and N. Jindal. An overview of the transmission capacity of wireless networks. *IEEE Trans. Commun.*, 58(12):3593 –3604, December 2010.

[73] R.K. Ganti, J.G. Andrews, and M. Haenggi. High-SIR transmission capacity of wireless networks with general fading and node distribution. *IEEE Trans. Inf. Theory*, 57(5):3100–3116, May 2011.

[74] G.J. Foschini and M.J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, 6:311–335, 1998.

[75] G.J. Foschini, G.D. Golden, R.A. Valenzuela, and P.W. Wolniansky. Simplified processing for high spectral efficiency wireless communication employing multi-element arrays. *IEEE J. Sel. Areas Commun.*, pages 1841–1852, 1999.

[76] A. Ghosh, J. Zhang, J. Andrews, and R. Muhamed. *Fundamentals of LTE*. Prentice Hall, 2010.

[77] B. Nosrat-Makouei, J.G. Andrews, and R.W. Heath. User arrival in MIMO interference alignment networks. *IEEE Trans. Wireless Commun.*, 11(2):842–851, February 2012.

[78] H. Zhai, Y. Kwon, and Y. Fang. Performance analysis of IEEE 802.11 MAC protocols in wireless LANs. *Wireless Communications and Mobile Computing*, 4(8):917–931, 2004.

[79] S.G. Sitharaman. Modeling queues using poisson approximation in IEEE 802.11 ad-hoc networks. In *The 14th IEEE Workshop on Local and Metropolitan Area Networks*, pages 6 pp.–6, September 2005.

[80] O. Tickoo and B. Sikdar. Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks. In *Proc. Twenty-third Annual Joint Conference of IEEE Computer and Communications Societies, INFOCOM*, volume 2, pages 1404–1413, March 2004.

[81] D. Stoyan, W. Kendall, and J. Mecke. *Stochastic geometry and its application*. Wiley, 2nd edition, 1996.

[82] L. Takács. *Introduction to the theory of queues*. Oxford University Press, 1962.

[83] C. Lee and M. Haenggi. Interference and outage in poisson cognitive networks. *IEEE Trans. Wireless Commun.*, 11(4):1392–1401, April 2012.

[84] Z. Gong and M. Haenggi. The local delay in mobile Poisson networks. *IEEE Trans. Wireless Commun.*, 12(9):4766–4777, September 2013.

[85] R.K. Ganti and M. Haenggi. Interference and outage in clustered wireless ad hoc networks. *IEEE Trans. Inf. Theory*, 55(9):4067–4086, September 2009.

[86] R. Tresch and M. Guillaud. Performance of interference alignment in clustered wireless ad hoc networks. In *Proc. IEEE International Symposium on Information Theory, ISIT*, pages 1703–1707, June 2010.

[87] R. Tresch, G. Alfano, and M. Guillaud. Interference alignment in clustered ad hoc networks: high reliablity regime and per-cluster ALOHA. In *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, pages 3348–3351, May 2011.

[88] B. Nosrat-Makouei, R.K. Ganti, J.G. Andrews, and R.W. Heath. Mimo interference alignment in random access networks. *IEEE Trans. Commun.*, 61(12):5042–5055, December 2013.

[89] Y. Zhong and W. Zhang. Multi-channel hybrid access femtocells: A stochastic geometric analysis. *IEEE Trans. Commun.*, 61(7):3016–3026, July 2013.

[90] W.C. Cheung, T.Q.S. Quek, and M. Kountouris. Throughput optimization, spectrum allocation, and access control in two-tier femtocell networks. *IEEE J. Sel. Areas Commun.*, 30(3):561–574, April 2012.

[91] J.G. Andrews, F. Baccelli, and R.K. Ganti. A tractable approach to coverage and rate in cellular networks. *IEEE Trans. Commun.*, 59(11):3122–3134, November 2011.

[92] T.X. Brown. Cellular performance bounds via shotgun cellular systems. *IEEE J. Sel. Areas Commun.*, 18(11):2443–2455, November 2000.

[93] R. Madan, J. Borran, Ashwin Sampath, N. Bhushan, A. Khandekar, and Tingfang Ji. Cell association and interference coordination in heterogeneous LTE-A cellular networks. *IEEE J. Sel. Areas Commun.*, 28(9):1479–1489, December 2010.

[94] M. Haenggi. On distances in uniformly random networks. *IEEE Trans. Inf. Theory*, 51(10):3584–3586, October 2005.

[95] H. ElSawy, E. Hossain, and M. Haenggi. Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey. *IEEE Communications Surveys Tutorials*, 15(3):996–1019, Third Quarter 2013.